

FIGURE 1

CCAATCGCCCGGTGCGGTGGTGCAGGGTCTCGGGCTAGTCA**ATG**GCGTCCCCGTCTCGGAGAC
TGCAGACTAAACCAGTCATTACTTGTTCAGAGCGTTCTGCTAATCTACACTTTTATTTTC
TGGATCACTGGCGTTATCCTTCTTGCAGTTGGCATTGTTGGGGCAAGGTGAGCCTGGAGAATTA
CTTTTCTCTTTTAAATGAGAAGGCCACCAATGTCCCCTTCGTGCTCATTGCTACTGGTACCG
TCATTATTCTTTTGGGCACCTTTGGTTGTTTTGCTACCTGCCGAGCTTCTGCATGGATGCTA
AAACTGTATGCAATGTTTCTGACTCTCGTTTTTTTTGGTCGAACTGGTCGCTGCCATCGTAGG
ATTTGTTTTTCAGACATGAGATTAAGAACAGCTTTAAGAATAATTATGAGAAGGCTTTGAAGC
AGTATAACTCTACAGGAGATTATAGAAGCCATGCAGTAGACAAGATCCAAAATACGTTGCAT
TGTTGTGGTGTCAACGATTATAGAGATTGGACAGATACTAATTATTACTCAGAAAAAGGATT
TCCTAAGAGTTGCTGTAAACTTGAAGATTGTACTCCACAGAGAGATGCAGACAAAGTAAACA
ATGAAGGTTGTTTTATAAAGGTGATGACCATTATAGAGTCAGAAATGGGAGTCGTTGCAGGA
ATTTCTTTTGGAGTTGCTTGCTTCCAACGATTGGAATCTTTCTCGCCTACTGCCWCTCTCG
TGCCATAACAAATAACCAGTATGAGATAGTG**TAA**CCCAATGTATCTGTGGGCCTATTCCTCT
CTACCTTTAAGGACATTTAGGGTCCCCCTGTGAATTAGAAAGTTGCTTGGCTGGAGAAGT
ACAACACTACTTACTGATAGACCAAAAACTACACCAGTAGGTTGATTCAATCAAGATGTAT
GTAGACCTAAAACTACACCAATAGGCTGATTCAATCAAGATCCGTGCTCGCAGTGGGCTGAT
TCAATCAAGATGTATGTTTGCTATGTTCTAAGTCCACCTTCTATCCCATTCTATGTTAGATCG
TTGAAACCCTGTATCCCTCTGAAACACTGGAAGAGCTAGTAAATTGTAAATGAAGT

FIGURE 2

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA19902
><subunit 1 of 1, 245 aa, 1 stop, 1 unknown
><MW: -1, pI: 8.36, NX(S/T): 1
MASPSRRLQTKPVITCFKSVLLIYTFIFWITGVILLAVGIWGKVSLENYFSLLNEKATNVPF
VLIATGTVIILLGTFGCFATCRASAWMLKLYAMFLTLVFLVELVAAIVGFVFRHEIKNSFKN
NYEKALKQYNSTGDIYRSHAVDKIQNTLHCCGVTDYRDWTDNYYSEKGFPSCKLEDCTPQ
RDADKVNNEGCFIKVMTIIESEMGVVAGISFGVACFQLIGIFLAYCXSRITNNQYEIV
```

Important features of the protein:**Signal peptide:**

amino acids 1-42

Transmembrane domains:

amino acids 19-42, 61-83, 92-114, 209-230,

N-glycosylation site.

amino acids 134-138

Tyrosine kinase phosphorylation site.

amino acids 160-168, 160-169

N-myristoylation site.

amino acids 75-81, 78-84, 210-216, 214-220, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 69-80, 211-222

FIGURE 3

CCCACGCGTCCGGCGCCGTGGCCTCGCGTCCATCTTTGCCGTTCTCTCGGACCTGTCACAAA
GGAGTCGCGCCGCCGCCGCCGCCCTCCCTCCGGTGGGCCCGGGAGGTAGAGAAAGTCAGT
GCCACAGCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCTGAGGGT
TGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACCGGACTGGG
GCCGGGGTAGGCTCTGGAAAGGGCCCCGGGAGAGAGGTGGCGTTGGTCAGAACCTGAGAAACA
GCCGAGAGGTTTTCCACCGAGGCCCGCGCTTGAGGGATCTGAAGAGGTTCCCTAGAAGAGGGT
GTTCCCTCTTTCCGGGGTCTCACCAGAAGAGGTTCTTGGGGTCCGCCCTTCTGAGGAGGCT
GCGGCTAACAGGGCCCAGAACTGCCATTGGATGTCCAGAATCCCCTGTAGTTGATAATGTTG
GGAATAAGCTCTGCAACTTTCTTTGGCATTTCAGTTGTTAAAAACAAATAGGATGCAAATTCC
TCAACTCCAGTTTATGAAAACAGTACTTGGAAGAACTGAAAACCTACCTAAATGATCGTCTTTG
GTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGGCCAGGGTCTGTTGTTGACTCTCGAAGAG
CACATAGCCCACTTCCTAGGGACTGGAGGTGCCGCTACTACCATGGGTAATTCCTGTATCTG
CCGAGATGACAGTGGAACAGATGACAGTGTGACACCCAACAGCAACAGGCCGAGAACAGTG
CAGTACCCACTGCTGACACAAGGAGCCAACCACGGGACCCTGTTCCGGCCACCAAGGAGGGGC
CGAGGACCTCATGAGCCAAGGAGAAAGAAACAAAATGTGGATGGGCTAGTGTTGGACACACT
GGCAGTAATACGGA CTCTTGTAGATAAGTAAGTATCTGACTCACGGTCACCTCCAGTGGAAT
GAAAAGTGTTCTGCCCCGAACCATGACTTTAGGACTCCTTCAGTTCCTTTAGGACATACTCG
CCAAGCCTTGTGCTCACAGGGCAAAGGAGAATATTTAATGCTCCGCTGATGGCAGAGTAAA
TGATAAGATTTGATGTTTTTGCTTGCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTC
TGTAGCAGAAAACACGATAAAGCTATGATCTTTATTAGAG

FIGURE 4

MIVFGWAVFLASRSLGQGLLLTLEEHIAHFLGTGGAATTMGNSCICRDDSGTDDSVDTQQQQ
AENSAVPTADTRSQPRDPVRPPRRGRGPHEPRRKKQNV DGLVLDTLAVIRTLVDKO

Signal peptide:

amino acids 1-16

Casein kinase II phosphorylation site.

amino acids 22-26, 50-54, 113-117

N-myristoylation site.

amino acids 18-24, 32-38, 34-40, 35-41, 51-57

FIGURE 5

GGCACGAGGCGCTGTCCACCCGGGGGCGTGGGAGTGAGGTACCAGATTCAGCCCATTGCGC
CCGACGCCTCTGTTCTCGGAATCCGGGTGCTGCGGATTGAGGTCCCGGTTCCCTAACGGACTG
CAAG**ATG**GAGGAAGGCGGGAACCTAGGAGGCCTGATTAAGATGGTCCATCTACTGGTCTTGT
CAGGTGCCTGGGGCATGCAAATGTGGGTGACCTTCGTCTCAGGCTTCCTGCTTTTCCGAAGC
CTTCCCCGACATACCTTCGGACTAGTGCAGAGCAAACCTCTTCCCCTTCTACTTCCACATCTC
CATGGGCTGTGCCTTCATCAACCTCTGCATCTTGGCTTCACAGCATGCTTGGGCTCAGCTCA
CATTCTGGGAGGCCAGCCAGCTTTACCTGCTGTTCCCTGAGCCTTACGCTGGCCACTGTCAAC
GCCCCGCTGGCTGGAACCCCGCACACAGCTGCCATGTGGGCCCTGCAAACCGTGGAGAAGGA
GCGAGGCCTGGGTGGGGAGGTACCAGGCAGCCACCAGGGTCCCGATCCCTACCGCCAGCTGC
GAGAGAAGGACCCCAAGTACAGTGCTCTCCGCCAGAATTTCTTCCGCTACCATGGGCTGTCC
TCTCTTTGCAATCTGGGCTGCGTCCTGAGCAATGGGCTCTGTCTCGCTGGCCTTGCCCTGGA
AATAAGGAGCCTC**TAG**CATGGGCCCTGCATGCTAATAAATGCTTCTTCAGAAATGAAAAAA
AAAAAAAAAAAA

FIGURE 6

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56107

<subunit 1 of 1, 231 aa, 1 stop

<NX(S/T): 0

MEEGGNLGGLIKMHLLVLSGAWGMQMWVTFVSGFLLFRSLPRHTFGLVQSKLFPFYFHISM
GCAFINLCILASQHAWAQLTFWEASQLYLLFSLTLATVNARWLEPRTTAAMWALQTVEKER
GLGGEVPGSHQGPDPYRQLREKDPKYSALRQNFFRYHGLSSLNLCVLSNGLCLAGLALEIRSL

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 86-103, 60-75

Casein kinase II phosphorylation site.

amino acids 82-86

Tyrosine kinase phosphorylation site.

amino acids 144-151

N-myristoylation site.

amino acids 4-10, 5-11, 47-53, 170-176, 176-182

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 54-65

G-protein coupled receptors proteins.

amino acids 44-85

FIGURE 7

AATTCAGATTTTAAAGCCATTCTGCAGTGGAATTCATGAACTAGCAAGAGGACACCATCTT
 CTTGTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGGAAAA**ATG**CTCTTTTGGGT
 GCTAGGCCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAG
 ACATCACTGATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCC
 AGAACTTTTGATAAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAAC
 AGCTTTAAAGGCAGAAACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAG
 AGAATGTCAAGAGGACTGCCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGT
 CTGATCAATAATGCTGGTGTTCCCGGCGTGCTGGCTCCCCTGACTGGCTGACACTAGAGGA
 CTACAGAGAACCTATTGAAGTGAACCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTC
 CTTTGGTCAAGAAAGCTCAAGGGAGAGTTATTAATGTCTCCAGTGTTGGAGGTCGCCTTGCA
 ATCGTTGGAGGGGGCTATACTCCATCCAAATATGCAGTGGAAGGTTTCAATGACAGCTTAAG
 ACGGGACATGAAAGCTTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAAACAA
 ACTTGGCAGATCCAGTAAAGGTAATTGAAAAAACTCGCCATTTGGGAGCAGCTGTCTCCA
 GACATCAACAACAATATGGAGAAGGTTACATTGAAAAAAGTCTAGACAACTGAAAGGCAA
 TAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTGCATGGACCACGCTCTAACAA
 GTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAATTTTCTGGATACCTCTG
 TCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGGCTAA
 TCCCAAGGCAGT**GTGA**CTCAGCTAACCACAAATGTCTCCTCCAGGCTATGAAATTGGCCGAT
 TTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACTCATTTAGA
 TCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGTCCCTG
 CTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCTGT
 ATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATG
 ATCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCAGCATTTACAGTAACTTGTGAATGTT
 AAGTATCATCTCTTATCTAAATATTTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56406
><subunit 1 of 1, 319 aa, 1 stop
><MW: 35227, pI: 8.97, NX(S/T): 3
MLFWVLGLLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKG FHVIAACLT
ESGSTALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDW
LTLEDYREPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGF
NDSLRRDMKAFGVHVSCIEPGLFKTNLADPVKVKIEKKLAIWEQLSPDIKQQYGEgyIEKSLD
KLKGNKSYVNMDLSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQK
AELANPKAV
```

Important features of the protein:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

CGCGGCTGTGACGGCGCTGCGGATGGCTGCCTGCGAGGGCAGGAGAAGCGGAGCTCTCGGTT
CCTCTCAGTCGGACTTCCTGACGCCGCCAGTGGGCGGGGCCCTTGGGCCGTGCGCCACCACT
GTAGTCATGTACCCACCGCCGCCGCCGCCCTCATCGGGACTTCATCTCGGTGACGCTGAG
CTTTGGCGAGAGCTATGACAACAGCAAGAGTTGGCGGCCGCGCTCGTGCTGGAGGAAATGGA
AGCAACTGTGAGATTGCAGCGGAATATGATTCTCTTCTCCTTGCCTTTCTGCTTTTCTGT
GGACTCCTCTTCTACATCAACTTGGCTGACCATTGGAAAGCTCTGGCTTTCAGGCTAGAGGA
AGAGCAGAAGATGAGGCCAGAAATTGCTGGGTTAAAACCAGCAAATCCACCCGTCTTACCAG
CTCCTCAGAAGGCGGACACCGACCCCTGAGAACTTACCTGAGATTTTCGTACACAGAAGACACA
AGACACATCCAGCGGGGACCACCTCACCTGCAGATTAGACCCCCAAGCCAAGACCTGAAGGA
TGGGACCCAGGAGGAGGCCACAAAAGGCAAGAAGCCCCCTGTGGATCCCCGCCCGGAAGGAG
ATCCGCAGAGGACAGTCATCAGCTGGAGGGGAGCGGTGATCGAGCCTGAGCAGGGCACCGAG
CTCCCTTCAAGAAGAGCAGAAAGTGGCCACCAAGCCTCCCCTGCCACCGGCCAGGACACAGGG
CACACCAGTGCATCTGAACTATCGCCAGAAGGGCGTGATTGACGTCTTCTTGCATGCATGGA
AAGGATACCGCAAGTTTGATGGGGCCATGACGAGCTGAAGCCTGTGTCCAGGTCTTTCAGT
GAGTGGTTTGGCCTCGGTCTCACACTGATCGACGCGCTGGACACCATGTGGATCTTGGGTCT
GAGGAAAGAATTTGAGGAAGCCAGGAAGTGGGTGTGGAAGAAGTTACACTTTGAAAAGGACG
TGGACGTCAACCTGTTTGAGAGCACGATCCGCATCCTGGGGGGGCTCCTGAGTGCCTACCAC
CTGTCTGGGGACAGCCTCTTCTGAGGAAAGCTGAGGATTTTGGAAATCGGCTAATGCCTGC
CTTCAGAACACCATCCAAGATTCTTACTCGGATGTGAACATCGGTACTGGAGTTGCCACAC
CGCCACGGTGGACCTCCGACAGCACTGTGGCCGAGGTGACCAGCATTACAGCTGGAGTTCGG
GAGCTCTCCCGTCTCACAGGGGATAAGAAGTTTCAGGAGGCAGTGGAGAAGGTGACACAGCA
CATCCACGGCCTGTCTGGGAAGAAGGATGGGCTGGTGCCCATGTTTCATCAATACCCACAGTG
GCCTCTTACCCACCTGGGCGTATTACGCTGGGCGCCAGGGCCGACAGCTACTATGAGTAC
CTGCTGAAGCAGTGGATCCAGGGCGGGAAGCAGGAGACACAGCTGCTGGAAGACTACGTGGA
AGCCATCGAGGGTGTGAGAACGCACCTGCTGCGGCACTCCGAGCCCAGTAAGCTCACCTTTG
TGGGGGAGCTTGCCACGGCCGCTTCAGTGCCAAGATGGACCACCTGGTGTGCTTCTTCCGCA
GGGACGCTGGCTCTGGGCGTCTACCACGGCCTGCCCGCCAGCCACATGGAGCTGGCCCAGGA
GCTCATGGAGACTTGTTACCAGATGAACCGGCAGATGGAGACGGGGCTGAGTCCCGAGATCG
TGCATTCAACCTTTACCCCCAGCCGGGCGCTCGGGACGTGGAGGTCAAGCCAGCAGACAGG
CACAACTGCTGCGGCCAGAGACCGTGGAGAGCCTGTTCTACCTGTACCGCGTCACAGGGGA
CCGCAAATACCAGGACTGGGGCTGGGAGATTCTGCAGAGCTTACGCCGATTACACGGGTCC
CCTCGGGTGGCTATTCTTCCATCAACAATGTCCAGGATCCTCAGAAGCCCGAGCCTAGGGAC
AAGATGGAGAGCTTCTTCTGGGGGAGACGCTCAAGTATCTGTTCTTGCTCTTCTCCGATGA
CCCAAACCTGCTCAGCCTGGACGCCTACGTGTTCAACACCGAAGCCACCCCTCTGCCTATCT
GGACCCCTGCCATAGGTTGGATGGCTGCTGGTGTGGGGACTTCGGGTGGGCAGAGGCACCTTG
CTGGGTCTGTGGCATTTTCCAAGGGCCCACGTAGCACCGGCAACCGCCAAGTGGCCCAGGCT
CTGAACTGGCTCTGGGCTCCTCCTCGTCTCTGCTTTAATCAGGACACCGTGAGGACAAGTGA
GGCCGTGAGTCTTGGTGTGATGCGGGGTGGGCTGGGCCGCTGGAGCCTCCGCCTGCTTCTC
CAGAAGACACGAATCATGACTCACGATTGCTGAAGCCTGAGCAGGTCTCTGTGGGCCGACCA
GAGGGGGGCTTCGAGGTGGTCCCTGGTACTGGGGTGACCGAGTGGACAGCCCAGGGTGCAGC
TCTGCCCAGGCTCGTGAAGCCTCAGATGTCCCAATCCAAGGTCTGGAGGGGCTGCCGTGA
CTCCAGAGGCCTGAGGCTCCAGGGCTGGCTCTGGTGTTTACAAGCTGGACTCAGGGATCCTC
CTGGCCGCCCCGCAGGGGGCTTGAGGGGCTGGACGGCAAGTCCGTCTAGCTCACGGGCCCCCT
CCAGTGGAATGGGTCTTTTCGGTGGAGATAAAAGTTGATTTGCTCTAACCGCAA

FIGURE 10

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56529

><subunit 1 of 1, 699 aa, 1 stop

><MW: 79553, pI: 7.83, NX(S/T): 0

MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVMYPPPPPPPHRDFISVTLSTFGESYDN
 SKSWRRRSCWRKWKQLSRLQRNMILFLAFLLCGLLFYINLADHWKALAFRLEEEQKMRPE
 IAGLKPANPPVLPAPQKADTDPENLPEISSQKTQRHIQRGPPHLQIRPPSQDLKDGTEQEEAT
 KRQEAPVDPRPEGDPQRTVISWRGAVIEPEQGTELPSSRAEVPTKPPLPPARTQGTPVHLNY
 RQKGVIDVFLHAWKGYRKFAWGHDELKPVSRSEFWFGLGLTLIDALDTMWILGLRKEFEEA
 RKWVSKKLHFEKDVDVNLFEISTIRILGGLLSAYHLSGDSLFLRKAEDFGNRLMPAFRTPSKI
 PYSDVNIGTGVAHPPRWTS DSTVAEVT SIQLEFRELSRLTGDKKFQEAVEKVTQHIHGLSGK
 KDGLVPMFINTHSGLETHLGVFTLGARADSYEYLLKQWIQGGKQETQLLEDYVEAIEGVRT
 HLLRHSEPSKLT FVGELAHGRFS AKMDHLVCFLPGTLALGVYHGLPASHMELAQELMETCYQ
 MNRQMETGLSPEIVHFNLYPQPGRRDVEVKPADRHLLRPETVESLFYLYRVTGDRKYQDWG
 WEILQSFSRFRTRVPSGGYSSINNVQDPQKPEPRDKMESFFLGETLKYLFLLFSDDPNLLSLD
 AYVFENTEAHPLPIWTPA

Important features of the protein:

Transmembrane domain:

amino acids 21-40 and 84-105 (type II)

FIGURE 11

GGCGCCGCGTAGGCCCGGGAGGCCGGGCCGGGCTGCGAGCGCCTGCCCCATGCGCCGC
CGCCTCTCCGCACG**ATG**TTCCCCTCGCGGAGGAAAGCGGCGCAGCTGCCCTGGGAGGACGGC
AGGTCCGGGTTGCTCTCCGGCGGCCTCCCTCGGAAGTGTTCCGTCTTCCACCTGTTTCGTGGC
CTGCCTCTCGCTGGGCTTCTTCTCCCTACTCTGGCTGCAGCTCAGCTGCTCTGGGGACGTGG
CCCGGGCAGTCAGGGGACAAGGGCAGGAGACCTCGGGCCCTCCCCGTGCCTGCCCCCAGAG
CCGCCCCCTGAGCACTGGGAAGAAGACGCATCCTGGGGCCCCCACC GCCTGGCAGTGCTGGT
GCCCTTCCGCGAACGCTTCGAGGAGCTCCTGGTCTTCGTGCCCCACATGCGCCGCTTCCTGA
GCAGGAAGAAGATCCGGCACCACATCTACGTGCTCAACCAGGTGGACCACTTCAGGTTCAAC
CGGGCAGCGCTCATCAACGTGGGCTTCCTGGAGAGCAGCAACAGCACGGACTACATTGCCAT
GCACGACGTTGACCTGCTCCCTCTCAACGAGGAGCTGGACTATGGCTTTCCTGAGGCTGGGC
CCTTCCACGTGGCCTCCCCGGAGCTCCACCCTCTCTACCACTACAAGACCTATGTCGGCGGC
ATCCTGCTGCTCTCCAAGCAGCACTACCGGCTGTGCAATGGGATGTCCAACCGCTTCTGGGG
CTGGGGCCGCGAGGACGACGAGTTCTACCGGCGCATTAAGGGAGCTGGGCTCCAGCTTTTCC
GCCCCTCGGGAATCACAACTGGGTACAAGACATTTGCCACCTGCATGACCCAGCCTGGCGG
AAGAGGGACCAGAAGCGCATCGCAGCTCAAAAACAGGAGCAGTTCAAGGTGGACAGGGAGGG
AGGCCTGAACACTGTGAAGTACCATGTGGCTTCCCGCACTGCCCTGTCTGTGGGCGGGGCCC
CCTGCACTGTCCTCAACATCATGTTGGACTGTGACAAGACCGCCACACCCTGGTGCACATTC
AGCT**TGA**GCTGGATGGACAGTGAGGAAGCCTGTACCTACAGGCCATATTGCTCAGGCTCAGGA
CAAGGCCTCAGGTTCGTGGGCCCAGCTCTGACAGGATGTGGAGTGGCCAGGACCAAGACAGCA
AGCTACGCAATTGCAGCCACCCGGCCGCCAAGGCAGGCTTGGGCTGGGCCAGGACACGTGGG
GTGCCTGGGACGCTGCTTGCCATGCACAGTGATCAGAGAGAGGCTGGGGTGTGTCTGTCCG
GGACCCCCCTGCCTTCCTGCTCACCTACTCTGACCTCCTTCACGTGCCCAGGCCTGTGGG
TAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCTACTCTGACCTCCTTCACGTGCCC
AGGCCTGTGGGTAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCCCCAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 12

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56531
><subunit 1 of 1, 327 aa, 1 stop
><MW: 37406, pI: 9.30, NX(S/T): 1
MFPSRRKAAQLPWEDGRSGLLSGGLPRKCSVFHLFVACLSLGGFFSLLWLQLSCSGDVARAVR
GQGQETSGPPRACPPEPPPEHWEEDASWGPHRLAVLVPFRERFEELLVFVPHMRRFLSRKKI
RHHIYVLNQVDHFRFNRAALINVGFLSSNSTDYIAMHDVDLLPLNEELDYGFPEAGPFHVA
SPELHPLYHYKTYVGGILLLSKQHYRLCNGMSNRFWGWGREDDFYRRIKGAGLQLFRPSGI
TTGYKTFRHLHDPAWRKRDQKRIAAQKQEQFKVDREGGLNTVKYHVASRTALSVGGAPCTVL
NIMLDCDKTATPWCTFS
```

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 29-49 (type II)

N-glycosylation site.

amino acids 154-158

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 27-31

Tyrosine kinase phosphorylation site.

amino acids 226-233

N-myristoylation site.

amino acids 19-25, 65-71, 247-253, 285-291, 303-309, 304-310

FIGURE 13

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCT**ATG**CTGCTGCTAACGCTGCTGCT
GCTGCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCT
CTTCTGCCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTC
CTGCCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGT
GGGGGATGGC**TAA**GAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGG
GCTGCCTTATTTAAAGTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGC
CCTGTTCAATTAAGAAATTGTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGT
AAATATGTCTTTATAATAAACAGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 14

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56862

<subunit 1 of 1, 73 aa, 1 stop

<MW: 7879, pI: 7.21, NX(S/T): 0

MLLLTLLLLLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFPPSFLCLLPHRPAMTCSQAQ
PRGEGEKVGDG

Important features:**Signal peptide:**

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

GGGACCCATGCGGCCGTGACCCCCCGGCTCCCTAGAGGCCAGCGCAGCCGACAGCGGACAAAG
GAGCATGTCCGCGCCGGGAAGGCCCGTCTCTCCGGCCGCCATAAGGCTCCGGTCGCCCGCTGG
GCCCCGCGCCGCGCTCCTGCCCCGCCGGGCTCCGGGGCGGCCCGCTAGGCCAGTGCGCCGCCG
CTCGCCCCGACAGGCCCGGCCCGCAGC**ATG**GAGCCACCCGGACGCCGGCGGGGCCGCGCGCA
GCCGCCGCTGTTGCTGCCGCTCTCGCTGTTAGCGCTGCTCGCGCTGCTGGGAGGCGGCGGGCG
GCGGCGGCGCCGCGGCGCTGCCCGCCGGCTGCAAGCACGATGGGCGGCCCCGAGGGGCTGGC
AGGGCGGCGGGCGCCGCCGAGGGCAAGGTGGTGTGCAGCAGCCTGGAACCTCGCGCAGGTCTT
GCCCCCAGATACTCTGCCCAACCGCACGGTCAACCCTGATTCTGAGTAACAATAAGATATCC
AGCTGAAGAATGGCTCATTTTTCTGGGTAAAGTCTCCTTGAAAGATTGGACCTCCGAAACAA
CTTATTAGTAGTATAGATCCAGGTGCCTTCTGGGGACTGTCATCTCTAAAAAGATTGGATCT
GACAAACAATCGAATAGGATGTCTGAATGCAGACATATTTGAGGACTCACCAATCTGGTTC
GGCTAAACCTTTGCGGGAATTTGTTTTCTTCATTATCTCAAGGAACTTTTGATTATCTTGCG
TCATTACGGTCTTTGGAATTCCAGACTGAGTATCTTTTGTGTGACTGTAACATACTGTGGAT
GCATCGCTGGGTAAAGGAGAAGAACATCACGGTACGGGATACCAGGTGTGTTTATCCTAAGT
CACTGCAGGCCCAACCAGTCAAGGCGTGAAGCAGGAGCTGTTGACATGCGACCCTCCGCTT
GAATTGCCGTCTTTCTACATGACTCCATCTCATCGCCAAGTTGTGTTTGAAGGAGACAGCCT
TCCTTTCCAGTGCATGGCTTCATATATTGATCAGGACATGCAAGTGTGTGGTATCAGGATG
GGAGAATAGTTGAAACCGATGAATCGCAAGGTATTTTTGTTGAAAAGAACATGATTCACAAC
TGCTCCTTGATTGCAAGTGCCCTAACCATTTCTAATATTACGGCTGGATCTACTGGAAATTG
GGGCTGTGTCATGTCCAGACCAAACGTGGGAATAATACGAGGACTGTGGATATTGTGGTATTAG
AGAGTTCTGCACAGTACTGTCCTCCAGAGAGGGTGGTAAACAACAAAGGTGACTTCAGATGG
CCCAGAACATTGGCAGGCATTACTGCATATCTGCAGTGTACGCGGAACACCCATGGCAGTGG
GATATATCCCGGAAACCCACAGGATGAGAGAAAAGCTTGCGCGCAGATGTGATAGAGGTGGCT
TTTGGGCAGATGATGATTATTCTCGCTGTCAGTATGCAATGATGTCACTAGAGTTCTTTAT
ATGTTTAATCAGATGCCCCCAATCTTACCAATGCCGTGGCAACAGCTCGACAGTTACTGGC
TTACACTGTGGAAGCAGCCAACTTTTTCTGACAAAATGGATGTTATATTTGTGGCAGAAATGA
TTGAAAAATTTTGAAGATTTACCAAGGAGGAAAAATCAAAAGAGCTAGGTGACGTGATGGTT
GACATTGCAAGTAACATCATGTTGGCTGATGAACGTGTCTGTGGCTGGCGCAGAGGGAAGC
TAAAGCCTGCAGTAGGATTGTGCAGTGTCTTCAGCGCATTGCTACCTACCGGCTAGCCGGTG
GAGCTCACGTTTATTCAACATATTCACCCAATATTGCTCTGGAAGCTTATGTCATCAAGTCT
ACTGGCTTCACGGGGATGACCTGTACCGTGTTCCAGAAAGTGGCAGCCTCTGATCGTACAGG
ACTTTTCGGATTATGGGAGGCGGGATCCAGAGGGAAACCTGGATAAGCAGCTGAGCTTTAAGT
GCAATGTTTCAAATACATTTTCGAGTCTGGCACTAAAGGTATGTTACATTCTGCAATCATTT
AAGACTATTTACAGTT**TAA**ATTAGAATGCTCCAAATGTTCTGCTTCGCAAATAACCTTATTA
AAAGATTTTTTTTTTGCAGGAAGATAGGTATTATTGCTTTTGCTACTGTTTTAAAGAAAATA
ACCAGGAAGAACTGCATTACGACTTTCAAGGGCCCTAGGCATTTTTGCCTTTGATTCCCTTT
CTTCACATAAAAAATATCAGAAATTACATTTTATAACTGCAGTGGTATAAATGCAATATACT
ATTGTTACATGTGAAAAAATTTTATTTGACTTAAAAGTTTATTTATTTGTTTTTTGCTCCT
GATTTTAAGACAATAAGATGTTTTCATGGGCCCCCTAAAAGTATCATGAGCCTTTGGCAGCTG
GCCTGCCAAGCCTAGTGGAGAAGTCAACCTGAGACCAGGTGTTTAAATCAAGCAAGCTGTAT
ATCAAAATTTTTTGGCAGAAAAACACAAATATGTCAATATCTTTTTTTAAAAAAGTATTTCA
TTGAAGCAAGCAAAATGAAAGCATTTTACTGATTTTTTAAATTTGGTGCTTTAGATATATTT
GACTACACTGTATTGAAGCAAATAGAGGAGGCACAACCTCCAGCACCCTAATGGAACCACATT
TTTTTCACTTAGCTTTTCTGTGGGCATGTGTAATTGTATTCTCTGCGGTTTTTAATCTCACAG
TACTTTATTTCTGTCTTGTCCCTCAATAATATCACAAACAATATTCCAGTCATTTTAATGGC
TGCATAATAACTGATCCAACAGGTGTTAGGTGTTCTGGTTTAGTGTGAGCACTCAATAAATA
TTGAATGAATGAACGAAAAAAAAAAAAAAAAA

FIGURE 16

MEPPGRRRRGRAQPPLLLPLSLLALLALLGGGGGGGAAALPAGCKHDGRPRGAGRAAGAAEGK
 VVCSSLELAQVLPDTPNRTVTLLILSNNKISELKNGSFSGLSLLERLDLRNNLISSIDPGA
 FWGLSSSLKRLDLTNNRIGCLNADIFRGLTNLVRLNLSGNLFSSLSQGTFDYLASLRSLFQT
 EYLLCDCNILWMHRWVKEKNITVRDTRCVYPKSLQAQPVTGVKQELLTCDPPELELPSFYMTF
 SHRQVVFEGLSLPFQCMASYIDQDMQVLWYQDGRIVETDESQGI FVEKNMIHNCSLIASALT
 ISNIQAGSTGNWGCHVQTKRGNNTRTVDIVVLESSAQYCPPEVVNNKGDFRWPRTLAGITA
 YLQCTRNTHGSGIYPGNPQDERKAWRRCDRGGFWADDDYSRCQYANDVTRVLYMFNQMPNLN
 TNAVATARQLLAYTVEAANFSDKMDVIFVAEMIEKFGRFTKEEKSKELGDMVDIASNIMLA
 DERVLWLAQREAKACSRIVQCLQRIATYRLAGGAHVYSTYSPNIALEAYVIKSTGFTGMTCT
 VFQKVAASDRTGLSDYGRRDPEGNLDKQLSFKCNVSNTFSSLALKVCYILQSFKTIYS

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 13-40 (type II)

N-glycosylation site.

amino acids 81-85, 98-102, 159-163, 206-210, 301-305, 332-336,
 433-437, 453-457, 592-596

N-myristoylation site.

amino acids 29-35, 30-36, 31-37, 32-38, 33-39, 34-40, 51-57,
 57-63, 99-105, 123-129, 142-148, 162-168, 317-323, 320-326,
 384-390, 403-409, 554-560

CCGTGGGGATGTCTAGGAGCTCGAAGGTGGTGGTGGGCCTCTCGGTGCTGCTGACGGCGGCC
ACAGTGGCCGGCGTACATGTGAAGCAGCAGTGGGACCAGCAGAGGGCTTCGTGACGGAGTTAT
CAGAGACATTGAGAGGCCAAATTCGGAAAAAAGAAAACATTCGTCTTTTGGGAGAACAGATTA
TTTTGACTGAGCAACTTGAAGCAGAAAGAGAGAAGATGTTATTGGCAAAGGATCTCAAAAA
TCATTGACTTGAATGTGAAATATCTGTTGGACAGACAACACGAGTTTGTGTGTGTGTGTTGAT
GGAGAGTAGCTTAGTAGTATCTTCATCTTTTTTTTTTGGTCACTGTCCTTTTAAACTTGATCA
AATAAAGGACAGTGGGTCATATAAGTTACTGCTTTCAGGGTCCCTTATATCTGAATAAAGGA
GTGTGGGCAGACACTTTTTTGAAGAGTCTGTCTGGGTGATCCTGGTAGAAGCCCCATTAGGG
TCACTGTCCAGTGCTTAGGGTTGTTACTGAGAAGCACTGCCGAGCTTGTGAGAAGGAAGGGA
TGGATAGTAGCATCCACCTGAGTAGTCTGATCAGTCGGCATGATGACGAAGCCACGAGAACA
TCGACCTCAGAAGGACTGGAGGAAGGTGAAGTGGAGGGAGAGACGCTCCTGATCGTCGAATCC

FIGURE 18

MSRSSKVVLGLSVLLTAATVAGVHVKQQWDQQRLRDGVIRDIERQIRKKENIRLLGEQIILT
EQLEAEREKMLLAKGSQKS

Signal peptide:

amino acids 1-21

18/249

FIGURE 19

CTGTCGTCTTTGCTTCAGCCGCAGTCGCCACTGGCTGCCTGAGGTGCTCTTACAGCCTGTTC
 CAAGTGTGGCTTAATCCGTCTCCACCACCAGATCTTTCTCCGTGGATTCTCTGCTAAGACC
 GCTGCCATGCCAGTGACGGTAACCCGCACCACCATCACAACCACCACGACGTCATCTTCGGG
 CCTGGGGTCCCCCATGATCGTGGGGTCCCCTCGGGCCCTGACACAGCCCCTGGGTCTCCTTCGC
 CTGCTGCAGCTGGTGTCTACCTGCGTGGCCTTCTCGCTGGTGGCTAGCGTGGGCGCCTGGAC
 GGGGTCCATGGGCAACTGGTCCATGTTACCTGGTGCTTCTGCTTCTCCGTGACCCTGATCA
 TCCTCATCGTGGAGCTGTGCGGGCTCCAGGCCCGCTTCCCCCTGTCTTGCGCAACTTCCCC
 ATCACCTTCGCCTGCTATGCGGCCCTCTTCTGCCTCTCGGCCTCCATCATCTACCCCACCAC
 CTATGTCCAGTTCCTGTCCCACGGCCGTTTCGCGGGACCACGCCATCGCCGCCACCTTCTTCT
 CCTGCATCGCGTGTGTGGCTTACGCCACCGAAGTGGCCTGGACCCGGGCCCCGGCCCCGGCGAG
 ATCACTGGCTATATGGCCACCGTACCCGGGCTGCTGAAGGTGCTGGAGACCTTCGTTGCCTG
 CATCATCTTCGCGTTCATCAGCGACCCCAACCTGTACCAGCACCAGCCGGCCCTGGAGTGGT
 GCGTGGCGGTGTACGCCATCTGCTTCATCCTAGCGGCCATCGCCATCCTGCTGAACCTGGGG
 GAGTGCACCAACGTGCTACCCATCCCCCTTCCCCAGCTTCTGTGCGGGGCTGGCCTTGCTGTC
 TGTCCCTCCTCTATGCCACCGCCCTTGTTCTCTGGCCCCCTCTACCAGTTCGATGAGAAGTATG
 GCGGCCAGCCTCGGCGCTCGAGAGATGTAAGCTGCAGCCGCAGCCATGCCTACTACGTGTGT
 GCCTGGGACCGCCGACTGGCTGTGGCCATCCTGACGGCCATCAACCTACTGGCGTATGTGGC
 TGACCTGGTGCACCTCTGCCCACCTGGTTTTTGTCAAGGTCTTAAGACTCTCCCAAGAGGCTCC
 CGTTCCCTCTCCAACCTCTTTGTTCTTCTTGCCCGAGTTTTCTTTATGGAGTACTTCTTTCC
 TCCGCCTTTCTCTGTTTTCTCTTCTGTCTCCCCCTCCCTCCCACCTTTTTCTTTCTTCC
 CAATTCTTGCACTCTAACCAGTTCTTGATGCATCTTCTTCTTCCCTTTCTCTTGCTGT
 TTCCTTCTGTGTTGTTTTGTTGCCACATCCTGTTTTACCCCTGAGCTGTTTCTCTTTTT
 CTTTTCTTTCTTTTTTTTTTTTTTTTTTAAAGACGGATTCTCACTCTGTGGCCAGGCTGGAG
 TGCAGTGGTGCATCTCAGCTCACTGCAACCCCCGCCTCCTGGGTCAAGCGATTCTCCTCC
 CCCAGCTCCCAAGTAGCTGGGAGGACAGGTGTGAGCTGCCGCACCCAGCCTGTTTCTCTTT
 TTCCACTCTTCTTTTTTCTCATCTCTTTTCTGGGTGCTGTGCGCTTCTTATCTGCCTGT
 TTTGCAAGCACCTTCTCCTGTGTCTTGGGAGCCCTGAGACTTCTTCTCTCCTTGCCCTCA
 CCCACCTCCAAAGGTGCTGAGCTCACATCCACACCCTTGAGCCGTCCATGCCACAGCCCC
 CCAAGGGGCCCCATTGCCAAAGCATGCCTGCCACCCTCGCTGTGCCTTAGTCAGTGTGTAC
 GTGTGTGTGTGTGTGTTTGGGGGGTGGGGGGTGGGTAGCTGGGGATTGGGCCCTCTTTCT
 CCCAGTGGAGGAAGGTGTGAGTGTACTTCCCCTTTAAATTAAAAACATATATATATATAT
 ATTTGGAGGTCAGTAATTTCCAATGGGCGGGAGGCATTAAGCACCGACCCTGGGTCCCTAGG
 CCCCCTGGCACTCAGCCTTGCCAGAGATTGGCTCCAGAATTTTGGCAGGCTTACAGAACAC
 CCACTGCCTAGAGGCCATCTTAAAGGAAGCAGGGGCTGGATGCCTTTCATCCCAACTATTCT
 CTGTGGTATGAAAAAG

FIGURE 20

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58727

<subunit 1 of 1, 322 aa, 1 stop

<MW: 35274, pI: 8.57, NX(S/T): 1

MPVTVTRTTITTTTTSSSGLGSPMIVGSPRALTQPLGLLRLLQLVSTCVAFSLVASVGAWTG
SMGNWSMFTWCFCFSVTLLIILIVELCGLQARFPLSWRNFPI TFACYAALFCLSASIIYPTTY
VQFLSHGRSRDHAI AATFFSCIACVAYATEVAWTRARPGEITGYMATVPGLLKVLET FVACI
IFAFISDPNLYQHQP ALEWCVAVYAICFILAAIAILLNLGECTNVLPIPFPSFLSGLALLSV
LLYATALVLWPLYQFDEKYGGQPRRSRDVSCSRSHAYYVCAWDRRLAVAILTAINLLAYVAD
LVHSAHLVFVKV

Important features:**Transmembrane domains:**

amino acids 41-60 (type II), 66-85, 101-120, 137-153, 171-192,
205-226, 235-255 and 294-312

N-glycosylation site.

amino acids 66-69

Glycosaminoglycan attachment site.

amino acids 18-21

FIGURE 21

FIGURE 22

MFLATLSFLLPFAHPFGTVSCEYMLGSPLSSLAQVNLSPFSPKPVHMDPNYCHPSTSLHLCS
 LAWSFTRLLHPPLSPGISQVVKDHVTKPTAMAQGRVAHLIEWKGWSKPSDSPAALESASFSSY
 SDLSEGEQEARFAAGVAEQFAIAEAKLRAWSSVDGEDSTDDSYDEDFAGGMDTDMAGQLPLG
 PHLQDLFTGHRFSRPVRQGSVEPESDCSQTVSPDTLCSSLCSLEDGLLGSPARLASQLLGDE
 LLLAKLPPSRESAFRSLGPLEAQDSLNSPLTESCLSPAEEEPAPCKDCQPLCPPLTGSWER
 QRQASDLASSGVVSLDEDEAEPEEQ

Signal peptide:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 123-127, 128-132, 155-159, 162-166, 166-170, 228-232,
 285-289, 324-328

Tyrosine kinase phosphorylation site.

amino acids 44-52

N-myristoylation site.

amino acids 17-23, 26-32, 173-179

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 11-22

FIGURE 23

GGTTCCTGGGCGCTCTGTTACACAAGCAAGATACAGCCAGCCCCACCTAATTTTGTTCCT
 GGCACCCCTCCTGCTCAGTGCACATTTGTCACACTTAACCCATCTGTTTTCTCTAATGCACGA
 CAGATTCCTTTTCAGACAGGACAACCTGTGATATTTTCAGTTCCTGATTGTAAATACCTCCTAAG
 CCTGAAGCTTCTGTTACTAGCCATTGTGAGCTTCAGTTTCTTCATCTGCAAAATGGGCATAA
 TACAATCTATTCTTGCCACATCAAGGGATTGTTATTCCTTTAAAAAAAACCAATACCAAAG
 AAGCCTACA**ATG**TTGGCCTTAGCCAAAATTCTGTTGATTTCAACGTTGTTTTATTCACTTCT
 ATCGGGGAGCCATGGAAAAGAAAATCAAGACATAAACACAACACAGAACATTGCAGAAGTTT
 TAAAAACAATGGAAAATAAACCTATTTCTTTGGAAAGTGAAGCAAACCTAAACTCAGATAAA
 GAAATATAACCACCTCAAATCTCAAGGCGAGTCATTCCCCTCCTTTGAATCTACCCAACAA
 CAGCCACGGAATAACAGATTTCTCCAGTAACTCATCAGCAGAGCATTCTTTGGGCAGTCTAA
 AACCCACATCTACCATTTCCACAAGCCCTCCCTTGATCCATAGCTTTGTTTCTAAAGTGCCT
 TGGAATGCACCTATAGCAGATGAAGATCTTTTGCCCATCTCAGCACATCCCAATGCTACACC
 TGCTCTGTCTTCAGAAAACCTTCACTTGGTCTTTGGTCAATGACACCGTGAAAACCTCCTGATA
 ACAGTTCATTACAGTTAGCATCCTCTCTTCAGAACCAACTTCTCCATCTGTGACCCCTTG
 ATAGTGGAACCAAGTGGATGGCTTACCACAAACAGTGATAGCTTCACTGGGTTTACCCCTTA
 TCAAGAAAAACAACCTCTACAGCCTACCTTAAAATTACCAATAATTCAAAACTCTTTCCAA
 ATACGTCAGATCCCCAAAAGAAAATAGAAATACAGGAATAGTATTTCGGGGCCATTTTAGGT
 GCTATTCTGGGTGTCTCATTGCTTACTCTTGTGGGCTACTTGTGTGTGGAAAAGGAAAAC
 GGATTCATTTTCCCATCGGCGACTTTATGACGACAGAAATGAACCAGTTCTGCGATTAGACA
 ATGCACCGGAACCTTATGATGTGAGTTTTGGGAATTCTAGCTACTACAATCCAACCTTTGAAT
 GATTCAGCCATGCCAGAAAGTGAAGAAAATGCACGTGATGGCATTCCCTATGGATGACATACC
 TCCACTTCGTACTTCTGTAT**TAG**AACTAACAGCAAAAAGGCGTTAAACAGCAAGTGTCTCTA
 CATCCTAGCCTTTTTGACAAATTCATCTTTCAAAGGTTACACAAAATTACTGTCACGTGGAT
 TTTGTCAAGGAGAATCATAAAAGCAGGAGACCAGTAGCAGAAATGTAGACAGGATGTATCAT
 CCAAAGGTTTTCTTTCTTACAATTTTTTGCCATCCTGAGGCATTTACTAAGTAGCCTTAATT
 TGTATTTTAGTAGTATTTTCTTAGTAGAAAATATTTGTGGAATCAGATAAAACTAAAAGATT
 TCACCATTACAGCCCTGCCTCATAACTAAATAATAAAAATTATTCCACCAAAAAATTCTAAA
 ACAATGAAGATGACTCTTTACTGCTCTGCCTGAAGCCCTAGTACCATAATTCAAGATTGCAT
 TTTCTTAAATGAAAATTGAAAGGGTGCTTTTTTAAAGAAAATTTGACTTAAAGCTAAAAAGAG
 GACATAGCCCAGAGTTTCTGTTATTGGGAAATTGAGGCAATAGAAATGACAGACCTGTATTC
 TAGTACGTTATAATTTTCTAGATCAGCACACACATGATCAGCCCACTGAGTTATGAAGCTGA
 CAATGACTGCATTCAACGGGGCCATGGCAGGAAAGCTGACCCTACCCAGGAAAGTAATAGCT
 TCTTTAAAGTCTTCAAAGGTTTTGGGAATTTTAACTTGTCTTAATATATCTTAGGCTTCAA
 TTATTTGGGTGCCTTAAAAACTCAATGAGAATCATGGT

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58732

><subunit 1 of 1, 334 aa, 1 stop

><MW: 36294, pI: 4.98, NX(S/T): 13

MLALAKILLISTLFYSLLSGSHGKENQDINTTQNI AEVFKTMENKPISLESEANLNSDKENI
TTSNLKASHSPPLNLPNNSHGITDFSSNSSAEHSLGSLKPTSTISTSPPLIHSEFVSKVPWNA
PIADEDLLPISAHPNATPALSSENFTWSLVNDTVKTPDNSSITVSILSSEPTSPSVTPLIVE
PSGWLTTNSDSFTGFTPYQEKTTLQPTLKFTNNSKLFNPNTSDPQKENRNTGIVFGAILGAIL
GVSLTLVGYLLCGKRKTD SFSHRRLYDDRNEPVLRLDNAPEPYDVSFGNSSYYNPTLNDSA
MPESEENARDGIPMDDIPPLRTSV

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 235-262

N-glycosylation site.

amino acids 30-34, 61-65, 79-83, 90-94, 148-152, 155-159,
163-167, 218-222, 225-229, 298-302, 307-311

FIGURE 25

AACAGGATCTCCTCTTGCAGTCTGCAGCCCAGGACGCTGATTCCAGCAGCGCCTTACCGCGC
AGCCCGAAGATTCACT**ATG**GTGAAAATCGCCTTCAATACCCCTACCGCCGTGCAAAAGGAGG
AGGCGCGGCAAGACGTGGAGGCCCTCCTGAGCCGCACGGTCAGAACTCAGATACTGACCGGC
AAGGAGCTCCGAGTTGCCACCCAGGAAAAAGAGGGGCTCCTCTGGGAGATGTATGCTTACTCT
CTTAGGCCTTTCATTCATCTTGGCAGGACTTATTGTTGGTGGAGCCTGCATTTACAAGTACT
TCATGCCCAAGAGCACCATTTACCGTGGAGAGATGTGCTTTTTTGATTCTGAGGATCCTGCA
AATTCCCTTCGTGGAGGAGAGCCTAACTTCCTGCCTGTGACTGAGGAGGCTGACATTCGTGA
GGATGACAACATTGCAATCATTGATGTGCCTGTCCCCAGTTTCTCTGATAGTGACCCCTGCAG
CAATTATTCATGACTTTGAAAAGGGAATGACTGCTTACCTGGACTTGTTGCTGGGGAACTGC
TATCTGATGCCCCCTCAATACTTCTATTGTTATGCCTCCAAAAAATCTGGTAGAGCTCTTTGG
CAAAGTGGCGAGTGGCAGATATCTGCCTCAAACCTTATGTGGTTCGAGAAGACCTAGTTGCTG
TGGAGGAAATTTCGTGATGTTAGTAACCTTGGCATCTTTATTTACCAACTTTGCAATAACAGA
AAGTCCTTCCGCCTTCGTGCGAGAGACCTCTTGCTGGGTTTCAACAAACGTGCCATTGATAA
ATGCTGGAAGATTAGACACTTCCCCAACGAATTTATTGTTGAGACCAAGATCTGTCAAGAG**T**
AAGAGGCAACAGATAGAGTGTCTTGGTAATAAGAAGTCAGAGATTTACAATATGACTTTAA
CATTAAAGGTTTATGGGATACTCAAGATATTTACTCATGCATTTACTCTATTGCTTATGCTTT
AAAAAAAGGAAAAAAAAAAAAAAAACTACTAACCCTGCAAGCTCTTGTCAAATTTTAGTTTAAT
TGGCATTGCTTGTTTTTTGAAACTGAAATTACATGAGTTTCATTTTTCTTTGCATTTATAG
GGTTTAGATTTCTGAAAGCAGCATGAATATATCACCTAACATCCTGACAATAAATTCATCC
GTTGTTTTTTTTGTTTGTGTTTTTTCTTTTCCTTTAAGTAAGCTCTTTATTCATCTTATG
GTGGAGCAATTTTAAATTTGAAATATTTTAAATTGTTTTTGAACTTTTTGTGTAAAATATA
TCAGATCTCAACATTGTTGGTTTCTTTTGTGTTTTTCATTTGTACAACCTTCTTGAATTTAGA
AATTACATCTTTGCAGTCTGTAGGTGCTCTGTAATTAACCTGACTTATATGTGAACAATT
TTCATGAGACAGTCATTTTAACTAATGCAGTGATTCTTCTCACTACTATCTGTATTGTGG
AATGCACAAAATTGTGTAGGTGCTGAATGCTGTAAGGAGTTTAGGTTGTATGAATTCTACAA
CCCTATAATAAATTTTACTCTATACAAAAAAAAAAAAAAAAAAAAA

FIGURE 26

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58828

<subunit 1 of 1, 263 aa, 1 stop

<MW: 29741, pI: 5.74, NX(S/T): 1

MVKIAFNTPTAVQKEEARQDVEALLSRTVRTQILTGKELRVATQEKEGSSGRCMLTLLGLSF
 ILAGLIVGGACIYKYFMPKSTIYRGEMCFFDSEDPANSLRGGEPNFLPVTEEADIREDDNIA
 IIDVPVPSFSDSDPAAIHDFEKGMTAYLDLLLGNCYLMPLNTSIVMPPKNLVELFGKLASG
 RYLPQTYVVREDLVAVEEIRDVSNLGIFIYQLCNNRKSFRLRRRDLLLGFNKRAIDKCWKIR
 HFPNEFIVETKICQE

Type II transmembrane domain:

amino acids 53-75

N-glycosylation site.

amino acids 166-170

Casein kinase II phosphorylation site.

amino acids 35-39, 132-136, 134-138

N-myristoylation site.

amino acids 66-72, 103-109

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 63-74

FIGURE 27

GGAGGAGGGAGGGCGGGCAGGCGCCAGCCCAGAGCAGCCCCGGGCACCAGCACGGACTCTCT
 CTTCCAGCCCAGGTGCCCCCACTCTCGCTCCATTTCGGCGGGAGCACCCAGTCCTGTACGCC
 AAGGAACTGGTCCTGGGGGCACC**ATG**GTTTCGGCGGCAGCCCCCAGCCTCCTCATCCTTCTG
 TTGCTGCTGCTGGGGTCTGTGCCTGCTACCGACGCCCCTGTGTGCCCTGAAGGCCACGTT
 CCTGGAGGATGTGGCGGGTAGTGGGGAGGCCGAGGGCTCGTCGGCCTCCTCCCCGAGCCTCC
 CGCCACCCTGGACCCCGGCCCTCAGCCCCACATCGATGGGGCCCCAGCCCACAACCCTGGGG
 GGCCCATCACCCCCCACCAACTTCCTGGATGGGATAGTGGACTTCTTCCGCCAGTACGTGAT
 GCTGATTGCTGTGGTGGGCTCCCTGGCCTTTCTGCTGATGTTTCATCGTCTGTGCCGCGGTCA
 TCACCCGGCAGAAGCAGAAGGCCTCGGCCTATTACCCATCGTCCTTCCCCAAGAAGAAGTAC
 GTGGACCAGAGTGACCGGGCCGGGGGCCCGGGCCTTCAGTGAGGTCCCCGACAGAGCCCC
 CGACAGCAGGCCCCGAGGAAGCCCTGGATTCTTCCCGGCAGCTCCAGGCCGACATCTTGGCCG
 CCACCCAGAACCTCAAGTCCCCCACCAGGGCTGCACTGGGCGGTGGGGACGGAGCCAGGATG
 GTGGAGGGCAGGGGCGCAGAGGAAGAGGAGAAGGGCAGCCAGGAGGGGGACCAGGAAGTCCA
 GGGACATGGGGTCCCAGTGGAGACACCAGAGGCGCAGGAGGAGCCGTGCTCAGGGGTCCTTG
 AGGGGGCTGTGGTGGCCGGTGAGGGCCAAGGGGAGCTGGAAGGGTCTCTCTTGTTAGCCCAG
 GAAGCCCAGGGACCAGTGGGTCCCCCGAAAGCCCCTGTGCTTGCAGCAGTGTCCACCCCAG
 TGTCT**TAA**CAGTCCTCCCGGGCTGCCAGCCCTGACTGTGCGGGCCCCCAAGTGGTCACCTCCCC
 GTGTATGAAAAGGCCTTCAGCCCTGACTGCTTCCTGACACTCCCTCCTTGGCCTCCCTGTGG
 TGCCAATCCCAGCATGTGCTGATTCTACAGCAGGCAGAAATGCTGGTCCCCGGTGCCCCGGA
 GGAATCTTACCAAGTGCCATCATCCTTCACCTCAGCAGCCCCAAAGGGCTACATCCTACAGC
 ACAGCTCCCCTGACAAAGTGAGGGAGGGCACGTGTCCCTGTGACAGCCAGGATAAAACATCC
 CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCCCAAACCTACTTTTTTAAACA
 GCTACAGGGTAAAATCCTGCAGCACCCACTCTGGAAAATACTGCTCTTAATTTTCCTGAAGG
 TGGCCCCCTGTTTCTAGTTGGTCCAGGATTAGGGATGTGGGGTATAGGGCATTAAATCCTC
 TCAAGCGCTCTCCAAGCACCCCCGGCCTGGGGGTGAGTTTCTCATCCCGCTACTGCTGCTGG
 GATCAGGTTGAATGAATGGAACCTTCCTGTCTGGCCTCCAAAGCAGCCTAGAAGCTGAGGG
 GCTGTGTTTGAGGGGACCTCCACCCTGGGGAAGTCCGAGGGGGCTGGGGAAGGGTTTCTGACG
 CCCAGCCTGGAGCAGGGGGGCCCTGGCCACCCCCCTGTTGCTCACACATTGTCTGGCAGCCTG
 TGTCCACAATATTCGTCACTCCTCGACAGGGAGCCTGGGCTCCGTCCCTGCTTTAGGGAGGCT
 CTGGCAGGAGGTCCTCTCCCCCATCCCTCCATCTGGGGCTCCCCCAACCTCTGCACAGCTCT
 CCAGGTGCTGAGATATAATGCACCAGCACATAAACCTTTATTCCGGCCTGAAAAAAAAAAAA
 AAAGA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58852

><subunit 1 of 1, 283 aa, 1 stop

><MW: 29191, pI: 4.52, NX(S/T): 0

MVSAAAPSLILLILLLLLSVLPATDARSVPLKATFLEDVAGSGEAEAGSSASSPSLPPPWTPAL
SPTSMGPQPTTLGGPSPPTNFLDGIVDFFRQYVMLIAVVGSLAFLLMFIVCAAVITRQKQKA
SAYYPSSFPPKKKYVDQSDRAGGPRAFSEVPDRAPDSRPEEALDSSRQLQADILAATQNLKSP
TRAAALGGGDGARMVEGRGAEEEEKGSQEGDQEVQGHGVPVETPEAQEEPCSGVLEGAVVAGE
GQGELEGSLLLAQEAQGPVGPPEPCACSSVHPSV

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 94-118

N-myristoylation site.

amino acids 18-24, 40-46, 46-52, 145-151, 192-198, 193-199,
211-217, 238-244, 242-248

FIGURE 29

201	201
202	202
203	203
204	204
205	205
206	206
207	207
208	208
209	209
210	210
211	211
212	212
213	213
214	214
215	215
216	216
217	217
218	218
219	219
220	220
221	221
222	222
223	223
224	224
225	225
226	226
227	227
228	228
229	229
230	230
231	231
232	232
233	233
234	234
235	235
236	236
237	237
238	238
239	239
240	240
241	241
242	242
243	243
244	244
245	245
246	246
247	247
248	248
249	249
250	250
251	251
252	252
253	253
254	254
255	255
256	256
257	257
258	258
259	259
260	260
261	261
262	262
263	263
264	264
265	265
266	266
267	267
268	268
269	269
270	270
271	271
272	272
273	273
274	274
275	275
276	276
277	277
278	278
279	279
280	280
281	281
282	282
283	283
284	284
285	285
286	286
287	287
288	288
289	289
290	290
291	291
292	292
293	293
294	294
295	295
296	296
297	297
298	298
299	299
300	300
301	301
302	302
303	303
304	304
305	305
306	306
307	307
308	308
309	309
310	310
311	311
312	312
313	313
314	314
315	315
316	316
317	317
318	318
319	319
320	320
321	321
322	322
323	323
324	324
325	325
326	326
327	327
328	328
329	329
330	330
331	331
332	332
333	333
334	334
335	335
336	336
337	337
338	338
339	339
340	340
341	341
342	342
343	343
344	344
345	345
346	346
347	347
348	348
349	349
350	350
351	351
352	352
353	353
354	354
355	355
356	356
357	357
358	358
359	359
360	360
361	361
362	362
363	363
364	364
365	365
366	366
367	367
368	368
369	369
370	370
371	371
372	372
373	373
374	374
375	375
376	376
377	377
378	378
379	379
380	380
381	381
382	382
383	383
384	384
385	385
386	386
387	387
388	388
389	389
390	390
391	391
392	392
393	393
394	394
395	3

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212

><subunit 1 of 1, 440 aa, 1 stop

><MW: 42208, pI: 6.36, NX(S/T): 1

MKFQGPLACLLALLCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
 AGSKVSEALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
 ADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGGLGGQGNPGGLGTPWVHGYPGNSAGSFGM
 NPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGS
 GSQSGSSGSGSNGDNNNGSSSGSSSGSSSGSSSGSSSGSSSGSSSGSSSGNSGGSRGDSGSESSW
 GSSTGSSSGNHGGSGGGNGHKGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRL
 LGGSGDNYRGQGSWSGGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQ
 RSSRIP

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
 74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
 155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
 218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
 249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
 279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
 295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
 323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
 383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 31

GACCGGTCCCTCCGGTCTTGGATGTGCGGACTCTGCTGCAGCGAGGGCTGCAGGCCCCGCGGGCGGTGCTCACCG
 TGCCCTGGCTGGTGGAGTTTCTCTCCTTTGCTGACCATGTTGTTCCCTTGCTGGAATATTACCGGGACATCTTCA
 CTCTCCTGCTGCGCCTGCACCGGAGCTTGGTGTGTGTCGAGGAGAGTGAGGGGAAGATGTTTCTTCTGAACAAGC
 TGCTGCTACTTGCTGTCTGGGCTGGCTTTTCCAGATTCCACAGTCCCTGAGGACTTGTCTTTCTGGAAGAGG
 GTCCCTCATATGCCTTTGAGGTGGACACAGTAGCCCCAGAGCATGGCTTGGACAATGCGCCTGTGGTGGACCAGC
 AGCTGCTCTACACCTGCTGCCCCCTACATCGGAGAGCTCCGGAAACTGCTCGCTTCGTGGGTGTCAGGCAGTAGTG
 GACGGAGTGGGGGCTTCATGAGGAAAATCACCCCCACCACTACCACCAGCCTGGGAGCCCAGCCTTCCCAGACCA
 GCCAGGGGCTGCAGGCACAGCTCGCCAGGCCCTTTTCCACAACCAGCCGCCCTCCTTGCGCCGACCGTAGAGT
 TCGTGGCAGAAAAGAAATTGGATCAAACCTGTGTCAAACATATCAAGGCTACACTGGTGGCAGATCTGGTGGCCAGG
 CAGAGTCACTTCTCCAAGAGCAGCTGGTGACACAGGGAGAGGAAGGGGGAGACCCAGCCCAGCTGTTGGAGATCT
 TGTGTTCCAGCTGTGCCCTCACGGGGCCCAGGCATTGGCCCTGGGGCGGGAGTTCTGTCAAAGGAAGAGCCCTG
 GGGCTGTGCGGGCGCTGCTTCCAGAGGAGACCCCGGCAGCCGTTCTGAGCAGTGCAGAGAACATTGCTGTGGGGC
 TTGCAACAGAGAAAGCCTGTGCTTGGCTGTGAGCCAACATCACAGCACTGATCAGGAGGGAGGTGAAAGCAGCAG
 TGAGTCGCACACTTCGAGCCCAGGGTCTGAACCTGCTGCCCGGGGGAGCGGAGGGGCTGCTCCCGCGCTGAGC
 GTGCTCTCCTTGCCCGTGGGGCCACGGGACCCCTGACGAGGGAGTCTCCCAGAGCATCTGGAACAGCTCCTAGGC
 CAGCTGGGCCAGACGCTGCGGTGCCGCCAGTTCTGTGCCACCTGCTGAGCAGCATCTGGCAAAGTGCTCTGTG
 GAGTTAGCTTCCCTCCTCGTTGCAGATCAAATTCCTATCCTAGGGCCCCCGGCACAGTACAGGCTGGAGAGAGGG
 CAGGCTCGAAGGCTTCTGCACATGCTGCTTTCTTGTGGAAGGAAGACTTTCAGGGGCCGTTCCGCTGCAGCTG
 CTGCTGAGCCCAAGAAATGTGGGGCTTCTGGCAGACACAAGGCCAAGGGAGTGGGACTTGCTGCTATTCTTGCTA
 CGGGAGCTGGTGGAGAAGGGTCTGATGGGACGGATGGAGATAGAGGCCTGCCTGGGCAGCCTCCACCAGGCCCCAG
 TGGCCAGGGGACTTTGCTGAAGAATTAGCAACACTGTCTAATCTGTTTCTAGCCGAGCCCCACCTGCCAGAACCC
 CAGCTAAGAGCCTGTGAGTTGGTGCAGCCAAACCGGGGCACTGTGCTGGCCCAGAGCTAGGGCTGAGAAGTGGCC
 CTGCCTTGGGCATTGCACCAGAACCCTGGACCCCCGCTCACGAGGAGGCCCAAGTGCCCAATGCAGACCCTCAC
 TGGTTGGGGTGTAGCTGGGTCTACAGTCAGACTTCTGCTCTAAGGGTGTCACTGCCTGGCATCCACCACGCGA
 ATCCTAGAGGAAGGAGAGTGGCCTGATTTGGGATTATGGCAGAAAAGTCCAGAGATGCCAGTCCCTGGAGTAGAA
 GAGGTGGTGTGTTGTTTATCTCTTGGATACTAAATGAAATGAGGTGTGTGGGCTTGTCAACACAGAATTCAAGCCT
 CATTTGCTATCCAGCATCTCTTAAACCTTTGTAGTCTTGGAATTCTGACAGAGGCAAAATGACTCCTGCTTAAC
 TTATGAAGAAAGTTAAACATGAATCTTGGGAGTCTACATTTTCTTATCACCAGGAGCTGGACTGCCATCTCCTT
 ATAAATGCCTAACACAGGCCCGGTCTGGTGGCTCATGCCTGTAATCCCAGCACTTTGAGAGGCTGAGGTGCGCG
 GACTGCCTGAGGTGAGGAATTCAAGACCAGCCTGGCCAACATGGCAAAACCCCATCTCTACTAAAAATAAAAAA
 TTATTAGCTGGGCATGGTGGTGTGTGCCTGTAATCCCAGCTACTCAGGAGGATGAGGCAGGAGACCTGCTTGAAC
 CTGGAGGTGGAGTTGCACTGAGCCGAGGTGCGACCACTGCACTCCAGTCTGGGTAAACAGAGCGAGACTTCTAG
 AAAAAAGCCTAACAAACAGATAAGGTAGGACTCAACCAACTGAAACCTGACTTCCCCCTGTACCTCAGCCCCCTG
 TGCAGGTAGTAACCTCTTGAGACCTCTCCCTGACCAGGGACCAAGCACAGGGCATTTAGAGCTTTTTTAGAATAAA
 CTGGTTTTCTTTAAAAAAGGGCTTTTATTAATAATTCTCCCCACACGATGGCTCCTGCAATCTGCCACAGCTC
 TGGGGCGTGTCTGTAGGGAAAGGCCCTGTTTTCCCTGAGGCGGGGCTGGGCTTGTCCATGGGTCCGCGGAGCTG
 GCCGTGCTTGGCGCCCTGGCGTGTGTCTAGCTGCTTCTTGCCGGGCACAGAGCTGCGGGGTCTGGGGGACCGGG
 AGCTAAGAGCAGGCTCTGGTGCAGGGGTGGAGGCCTGTCTCTTAACCGACACCCCTGAGGTGCTCCTGAGATGCTG
 GGTCCACCCTGAGTGGCACGGGGAGCAGCTGTGGCCGGTGTCTCTCYTAGGCCAGTCTCTGGGGAAACTAAGCTC
 GGGCCCTTCTTTGCAAAGACCGAGGATGGGGTGGGTGTGGGGGACTCATGGGGAATGGCCTGAGGAGCTACGTGT
 GAAGAGGGCGCCGTTTGTGGCTGCAGCGGCCTGGAGCGCCTCTCTCCTGAGCCTCAGTTTCCCTTTCCGTCTA
 ATGAAGAACATGCCGTCTCGGTGTCTCAGGGCTATTAGGACTTGGCCTCAGGAAGTGGCCTTGGACGAGCGTCAT
 GTTATTTTCACAACTGTCCTGCGACGTTGGCCTGGGCACGTGATGGAATGGCCCATGTCCCTCTGCTGCGTGGAC
 GTCGCGGTGCGGAGTGCAGCAGCCAGAGGCGGGGCCAGACGTGCGCCTGGGGGTGAGGGGAGGCGCCCCGGGAGGG
 CCTCACAGGAAGTTGGGCTCCCGCACCACCAGGCAGGGCGGGCTCCCGCCGCCGCCGCCACCACCGTCCAGG
 GGCCGGTAGACAAAGTGGAAGTCGCGCTTGGGCTCGCTGCGCAGCAGGTAGCCCTTGATGCAGTGGGCGAGCGG
 TCGTCCGCCAGCTGGAAGCAGCGCCCGTCCACCAGCACGAACAGCCGGTGGCCT

FIGURE 32

MCFLNKLALLLAVLGWLFQIPTVPEDLFFLEEGPSYAFEVDTVAPEHGLDNAPVVDQQLLYTC
CPYIGELRKLLASWVSGSSGRSGGFMRKITPTTTTSLGAQPSQTSQGLQAQLAQAFFHNQPP
SLRRTVEFVAERIGSNCVKHIKATLVADLVRQAESLLQEQLVTQGEEGGDPAQLLEILCSQL
CPHGAQALALGREFCQRKSPGAVRALLPEETPAAVLSSAENIAVGLATEKACAWLSANITAL
IRREVKA AVSRTLRAQGPEPAARGERRGCSRA

Signal peptide:

amino acids 1-18

N-glycosylation site.

amino acids 244-248

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 89-93

Casein kinase II phosphorylation site.

amino acids 21-25, 167-171, 223-227

N-myristoylation site.

amino acids 100-106, 172-178, 207-213

Microbodies C-terminal targeting signal.

amino acids 278-282

FIGURE 33

TCCCTTGACAGGTCTGGTGGCTGGTTTCGGGGTCTACTGAAGGCTGTCTTGATCAGGAACTG
 AAGACTCTCTGCTTTTGCCACAGCAGTTTCCTGCAGCTTCCTTGAGGTGTGAACCCACATCCC
 TGCCCCCAGGGCCACCTGCAGGACGCCGACACCTACCCCTCAGCAGACGCCGGAGAGAA**ATG**
 AGTAGCAACAAAGAGCAGCGGTGAGCAGTGTTCGTGATCCTCTTTGCCCTCATCACCATCCT
 CATCCTCTACAGCTCCAACAGTGCCAATGAGGTCTTCCATTACGGCTCCCTGCGGGGCCGTA
 GCCGCCGACCTGTCAACCTCAAGAAGTGGAGCATCACTGACGGCTATGTCCCCATTCTCGGC
 AACAAGACACTGCCCTCTCGGTGCCACCAAGTGTGTGATTGTCAGCAGCTCCAGCCACCTGCT
 GGGCACCAAGCTGGGCCCTGAGATCGAGCGGGCTGAGTGTACAATCCGCATGAATGATGCAC
 CCACCACTGGCTACTCAGCTGATGTGGGCAACAAGACCACCTACCGCGTCTGTGGCCCCATTCC
 AGTGTGTTCCGCGTGCTGAGGAGGCCCCAGGAGTTTGTCAACCGGACCCCTGAAACCGTGTT
 CATCTTCTGGGGGGCCCCGAGCAAGATGCAGAAGCCCCAGGGCAGCCTCGTGCGTGTGATCC
 AGCGAGCGGGCCTGGTGTTCCCCAACATGGAAGCATATGCCGTCTCTCCCGGCCGCATGCGG
 CAATTTGACGACCTCTTCCGGGGTGAGACGGGCAAGGACAGGGAGAAAGTCTCATTCGTGGTT
 GAGCACAGGCTGGTTTACCATGGTGATCGCGGTGGAGTTGTGTGACCACGTGCATGTCTATG
 GCATGGTCCCCCCCCAACTACTGCAGCCAGCGGCCCCGCCTCCAGCGCATGCCCTACCACTAC
 TACGAGCCCAAGGGGCCGGACGAATGTGTACCTACATCCAGAATGAGCACAGTCGCAAGGG
 CAACCACCACCGCTTCATCACCCGAGAAAAGGGTCTTCTCATCGTGGGCCCAGCTGTATGGCA
 TCACCTTCTCCCACCCCTCCTGGACCT**TAG**GCCACCCAGCCTGTGGGACCTCAGGAGGGTCAG
 AGGAGAAGCAGCCTCCGCCCAGCCGCTAGGCCAGGGACCATCTTCTGGCCAATCAAGGCTTG
 CTGGAGTGTCTCCAGCCAATCAGGGCCTTGAGGAGGATGTATCCTCCAGCCAATCAGGGCC
 TGGGGAATCTGTTGGCGAATCAGGGATTTGGGAGTCTATGTGGTTAATCAGGGGTGTCTTTC
 TTGTGCAGTCAGGGTCTGCGCACAGTCAATCAGGGTAGAGGGGGTATTTCTGAGTCAATCTG
 AGGCTAAGGACATGTCTTTCCCATGAGGCCTTGGTTTCAAGCCCCAGGAATGGACCCCCCA
 ATCACTCCCCACTCTGCTGGGATAATGGGGTCTGTCCCAAGGAGCTGGGAACCTTGGTGTTG
 CCCCCTCAATTTCCAGCACCAGAAAGAGAGATTGTGTGGGGGTAGAAGCTGTCTGGAGGCC
 GGCCAGAGAATTTGTGGGGTTGTGGAGGTTGTGGGGCGGTGGGGAGGTCCCAGAGGTGGGA
 GGCTGGCATCCAGGTCTTGGCTCTGCCCTGAGACCTTGGACAAACCCTTCCCCCTCTCTGGG
 CACCCTTCTGCCCCACACCAAGTTTCCAGTGCGGAGTCTGAGACCCTTTCCACCTCCCCTACAA
 GTGCCCTCGGGTCTGTCTCCCCGTCTGGACCCTCCCAGCCACTATCCCTTGCTGGAAGGCT
 CAGCTCTTTGGGGGGTCTGGGGTGACCTCCCCACCTCCTGGAAAACCTTTAGGGTATTTTTGC
 GCAAACCTCCTTCAGGGTTGGGGGACTCTGAAGGAAACGGGACAAAACCTTAAGCTGTTTTCT
 TAGCCCCCTCAGCCAGCTGCCATTAGCTTGGCTCTTAAAGGGCCAGGCCTCCTTTTCTGCCCT
 CTAGCAGGGAGGTTTTCCAACCTGTTGGAGGCGCCTTTGGGGCTGCCCTTTGTCTGGAGTCA
 CTGGGGGCTTCCGAGGGTCTCCCTCGACCCTCTGTCTGCTGGGATGGCTGTCTGGGAGCTGT
 ATCACCTGGGTTCTGTCCCCCTGGCTCTGTATCAGGCACTTTATTAAAGCTGGGCCTCAGTGG
 GGTGTGTTTGTCTCCTGCTCTTCTGGAGCCTGGAAGGAAAGGGCTTCAGGAGGAGGCTGTGA
 GGCTGGAGGGACCAGATGGAGGAGGCCAGCAGCTAGCCATTGCACACTGGGGTGATGGGTGG
 GGGCGGTGACTGCCCCAGACTTGGTTTTGTAATGATTTGTACAGGAATAAACACACCTACGC
 TCCGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 34

MSSNKEQRSVAVFVILFALITILILYSSNSANEVFHYGSLRGRSRRPVNLKKWSITDGYVPIL
 GNKTLPSRCHQCIVIVSSSSHLLGTLGPEIERAECTIRMNDAPTTGYSADVGNKTTYRVVAH
 SSVFRVLRRPQEFVNRTPETVFIFWGPPSKMQKPQGSLVRVIQRAGLVFPNMEAYAVSPGRM
 RQFDDLFRGETGKDREKSHSWLSTGWFTMVI AVELCDHVHVYGMVPPNYCSQRPRLQRMPIYH
 YYEPKGPDECVTYIQNEHSRKGNHHRFITEKRVFSSWAQLYGITFSHPSWT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 9-31 (type II)

N-glycosylation site.

amino acids 64-68, 115-119

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 50-54

Casein kinase II phosphorylation site.

amino acids 3-7, 29-33, 53-57, 197-201

Tyrosine kinase phosphorylation site.

amino acids 253-262

N-myristoylation site.

amino acids 37-43, 114-120, 290-294

FIGURE 35

GTTCTCATAGTTGGCGTCTTCTAAAGGAAAAACACTAAAATGAGGAACTCAGCGGACCGGGAGCGACGCAGCTT
 GAGGGAAGCATCCCTAGCTGTTGGCGCAGAGGGGCGAGGCTGAAGCCGAGTGGCCCGAGGTGTCTGAGGGGCTGG
 GGCAAAGGTGAAAGAGTTTTCAGAACAAAGCTTCTGGAACCCATGACCCATGAAGTCTTGTGACATTTATACCGT
 CTGAGGGTAGCAGCTCGAAACTAGAAGAAGTGGAGTGTGCCAGGGACGGCAGTATCTCTTTGTGTGACCCCTGGC
 GGCCTATGGGACGTTGGCTTCAGACCTTTGTGATACACC**ATG**CTGCGTGGGACGATGACGGCGTGGAGAGGAATG
 AGGCCTGAGGTACACTGGCTTGCCTCCTCCTAGCCACGAGGCTGCTTTGCTGACTTGAACGAGGTCCCTCAG
 GTCACCCGTCCAGCCTGCGTCCACCGTCCAGAAGCCCGGAGGCACTGTGATCTTGGGCTGCGTGGTGGAACTCCA
 AGGATGAATGTAACCTGGCGCCTGAATGGAAAGGAGCTGAATGGCTCGGATGATGCTCTGGGTGTCTCATCACC
 CACGGGACCCCTCGTCATCACTGCCCTTAACAACCACTGTGGGACGGTACCAGTGTGTGGCCCGGATGCCCTGCG
 GGGCTGTGGCCAGCGTGCCAGCCACTGTGACCTGACCTAGCCATCTCCAGGACTTCAAGTTAGATGTGCGACACGTG
 ATTGAAGTGGATGAGGGAAACACAGCAGTCATTGCCTGCCACCTGCCTGAGAGCCACCCCAAAGCCCAGGTCCGG
 TACAGCGTCAAACAAGAGTGGCTGGAGGCTCCAGAGGTAACCTGATCATGCCCTCAGGGAACCTCCAGATT
 GTGATGCCAGCCAGGAGGACGAGGGCATGTACAAGTGTGACGCTTACAACCCAGTGACCCAGGAAGTAAAAACC
 TCCGGCTCCAGCGACAGGTCAGTGTGCGCGCTCCAGCTGAGGCTGCCCGCATCATCTACCCCCGAGGGCC
 CAAACCATCATCGTCACCAAAGGCCAGAGTCTCATTTCTGGAGTGTGTGGCCAGTGGAAATCCACCCCCACGGGTC
 ACCTGGGCCAAGGATGGGTCCAGTGTACCGGCTACAACAAGACGCGCTTCTGCTGAGCAACCTCCTCATCGAC
 ACCACCGAGGAGGACTCAGGCACCTACCGCTGCATGGCCGACAATGGGGTTGGGCAGCCCGGGGACGGGTC
 ATAACCCCAAGGCTATGGCAGGATGCTGAGCTGGCTACCTGACGACACCTCCTGTATCACCTGCCAACTCCGGC
 CAGAGTGCCAAGCTTACCTGTGAGGTGCGTGGGAACCCCGCCCTCCGTGCTGTGGCTGAGGAATGCTGTGCCC
 CTCATCTCCAGCCAGCGCCTCCGGCTCTCCCGCAGGGCCCTGCGCGTGTCTCAGCATGGGGCTGAGGACGAAGGC
 GTCTACCAAGTGCATGGCCGAGAACGAGGTTGGGAGCGCCCATGCCGTAGTCCAGCTGCGGACCTCCAGGCCAAGC
 CGCAAGCAGGTACAAATTCTCTGACGATTGGACCATCTCTGGCATTCCAGCCAACACAGCACCGCCTGACCTC
 ACCAGACTTGACCCCGGAGCTTGTATGAAGTGGAGATGGCAGCTTACAACCTGTGCGGGAGAGGGCCAGACAGCC
 ATGGTCACCTTCCGAAGTGGACGGCGGCCCAAACCCGAGATCATGGCCAGCAAAGAGCAGCAGATCCAGAGAGAC
 GACCCTGGAGCCAGTCCCCAGAGCAGCAGCCAGCCAGACCACGGCCGCTCTCCCCCCAGAAGCTCCCGACAGG
 CCCACCATCTCCACGGCCTCCGAGACCTCAGTGTACGTGACCTGAGGATTCCCGTGGGAATGGTGGGTCCCAATC
 CAGTCTTCCGTGTGGAGTACAAGAAGCTAAAGAAAGTGGGAGACTGGATTCTGGCCACCAGCGCCATCCCCCA
 TCGCGGCTGTCCGTGGAGATCACGGGCTAGAGAAAGGCACCTCCTACAAGTTTCGAGTCCGGGCTCTGAACATG
 CTGGGGGAGAGCGAGCCAGCGCCCTCTCGGCCCTACGTGGTGTGCGGCTACAGCGGTCCGGTGTACGAGAGG
 CCGTGGCAGGTCTTATATACGCTTACGGTACGGTCAATGAGACCACCATCATGCTCAAGTGGATGACATC
 CCAGCAAGTAACAACAACACCCCAATCCATGGCTTTTATATCTATTATCGACCCACAGACAGTGACAATGATAGT
 GACTACAAGAAGGATATGGTGGAAAGGGGACAAGTACTGGCACTCCATCAGCCACCTGCAGCCAGAGACCTCCTAC
 GACATTAAGATGCAGTGCTTCAATGAAGGAGGGGAGAGCGAGTTCAGCAACGTGATGATCTGTGAGACCAAGCT
 CGGAAGTCTTCTGGCCAGCCTGGTGCAGTGCCACCCCCAACTCTGGCCCCACCACAGCCGCCCTTCTCTGAAACC
 ATAGAGCGGCCGGTGGGCACTGGGGCCATGGTGGCTCGCTCCAGCGACCTGCCCTATCTGATTGTGCGGGTCTGTC
 CTGGGCTCCATCGTTCTCATCATCGTCACCTTCATCCCCTTCTGCTTGTGGAGGGCCTGGTCTAAGCAAAAACAT
 ACAACAGACCTGGGTTTTCTCGAAGTGCCTTCCACCCCTCCTGCCCGTATACTATGGTGCCATTGGGAGGACTC
 CCAGGCCACACAGCCAGTGGACAGCCCTACCTCAGTGGCATCAGTGGACGGGCTGTGCTAATGGGATCCACATG
 AATAGGGGCTGCCCCCTCGGCTGCAGTGGGCTACCCGGGCTGAAGCCCCAGCAGCACTGCCAGGGCAGCTTCAG
 CAGCAGAGTGACACCAGCAGCCTGCTGAGGCAGACCCATCTTGGCAATGGATATGACCCCAAAGTCACCAGATC
 ACGAGGGGTCCCAAGTCTAGCCCGGACGAGGGCTCTTTCTTATACACTGCCCGACGACTCCACTCACCAGCTG
 CTGCAGCCCCATCAGCACTGCTGCCAACGCCAGGAGCAGCCTGCTGTGGGCCAGTCAGGGGTGAGGAGAGCC
 CCGACAGTCTGTCTGGAAGCAGTGTGGGACCTCCATTTCACTCAGGGCCCCCATGCTGCTTGGGCTTGTG
 CCAGTTGAAGAGGTGGACAGTCTGACTCCTGCCAAGTGAAGTGGAGGAGACTGGTGTCCCCAGCACCCCGTAGGG
 GCCTACGTAGGACAGGAACCTGGAATGCAGTCTCCCCGGGGCCACTGGTGGTGTGCTTTTTGAAACACCACCT
 CTCACAATTT**TAGG**CAGAAGCTGATATCCGAAAGACTATATATTGTTTTTTTTTAAAAAAGAGAAAAA
 AGAGACAGAGAAAATTGGTATTTATTTTTCTATTATAGCCATATTTATATATTTATGCACTTGTAATAAATGTA
 TATGTTTTATAATTCTGGAGAGACATAAGGAGTCTACCCGTTGAGGTTGGAGAGGGAAATAAAGAAGCTGCCA
 CCTAACAGGAGTCAACCAGGAAAGCACCGCAGAGGCTGGCGCGGGACAGACTCCTAACCTGGGGCCTCTGCAGTG
 GCAGGCGAGGCTGCAGGAGGGCCACAGATAAGCTGGCAAGAGGAAGGATCCAGGCACATGGTTCATCAGGAGCA
 TGAGGGAACAGCAAGGGGCACGGTATCACAGCCTGGAGACACCCACACAGATGGCTGGATCCGGTGTACGGGAA
 ACATTTTCTAAGATGCCCATGAGAACAGACCAAGATGTGTACAGCACTATGAGCATTAATAAACCTTCCAGAAT
 CAATAATCCGTGGCAACATATCTCTGTAAAAACAAACACTGTAACCTCTAAATAAATGTTTAGTCTTCCCTGTAAAA

FIGURE 36

MLRGTMTAWRGMREVTLACLLLATAGCFADLNEVPQVTVQPASTVQKPGGTVILGCVVEPP
 RMNVTWRLNGKELNGSDDALGLVLIHGTLVITALNNHTVGRYQCVARMPAGAVASVPATVTL
 ANLQDFKLDVQHVIEWDEGNTAVIACHLPESHHPKAQVRYSVKQEWLEASRGNYLIMPSGNLQ
 IVNASQEDEGMYKCAAYNPVTQEVKTSGSSDRLRVRRSTAEAAARI IYPPEAQTIIVTKGQSL
 ILECVASGI PPPRVTWAKDGSSVTGYNKTRFLLSNLLIDTTSEEDSGTYRCMADNGVGQPGA
 AVILYNVQVFEPPEVTMELSOLVIPWGQSAKLTCEVRGNPPPSVLWLRNAVPLISSQRLRLS
 RRALRVLSMGPEDEGVYQCMAENEVGSAAHAVVQLRTSRPSITPRLWQDAELATGTPPVSPSK
 LGNPEQMLRGQPALPRPPTSVGPASPKCPGEGKGQAPAEAPIILSSPRTSKTDSYELVWRPR
 HEGSGRAPILYYVVKHRKQVTNSSDDWTISGIPANQHRLTLTRLDPGSLYEVEMAAYNCAGE
 GQTAMVTFRTGRRPKPEIMASKEQQIQRDDPGAS PQSSSQPDHGRLSPPEAPDRPTISTASE
 TSVYVTWIPRNGGGFPIQSFVEYKKLKKVGDWILATSAIPPSRLSVEITGLEKGTSYKFRV
 RALNMLGESEPSAPSRPYVVS GYSGRVYERP VAGPYITFTDAVNETTIMLKWMIIPASNNNT
 PIHGFYIYYRPTDSDNDSYKKDMVEGDKYWHSISHLQPETS YDIKMQCFNEGGESEFSNVM
 ICETKARKSSGQPGRLP PPTLAPPQPPLPETIERPVGTGAMVARSSDLPYLIVGVVLGSIVL
 IIVTFIPFCLWRAW SKQKHTTDLGFPR SALPPSCPYTMVPLGG LPGHQASGQPYLSGISGRA
 CANGIHMNRGCPSAAVGYPGMKPQQHCPGELQQQSDTSSLLRQTHLGNGYDPQSHQITRGPK
 SSPDEGSFLYTL PDDSTHQLLQPHHDCCQRQE QPAAVGQSGVRRAPDSPVLEAVWDPPFHSG
 PPCCCLGLVPVEEVDSPDSCQVSGGDWCPQHVPV GAYVGQEPGMQLSPGPLVRVSFETPPLTI

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 16-30 (type II), 854-879

FIGURE 37

CGGGAGGCTGGGTCGTCATGATCCGGACCCCATTTGTCGGCCTCTGCCCATCGCCTGCTCCTC
 CCAGGCTCCCGCGGCCGACCCCCGCGCAAC**ATG**CAGCCACGGGCCGCGAGGGTTCCCGCGC
 GCTCAGCCGGCGGTATCTGCGGCGTCTGCTGCTCCTGCTACTGCTGCTGCTGCTGCGGCAGC
 CCGTAACCCGCGCGGAGACCACGCCGGGCGCCCCCAGAGCCCTCTCCACGCTGGGCTCCCCC
 AGCCTCTTCACCACGCCGGGTGTCCCCAGCGCCCTCACTACCCAGGCCTCACTACGCCAGG
 CACCCCCAAAACCCTGGACCTTCGGGGTCGCGCGCAGGCCCTGATGCGGAGTTTCCCACTCG
 TGGACGGCCACAATGACCTGCCCCAGGTCTGAGACAGCGTTACAAGAATGTGCTTCAGGAT
 GTTAACCTGCGAAATTTAGCCATGGTCAGACCAGCCTGGACAGGCTTAGAGACGGCCTCGT
 GGGTGCCAGTTCTGGTCAGCCTCCGTCTCATGCCAGTCCCAGGACCAGACTGCCGTGCGCC
 TCGCCCTGGAGCAGATTGACCTCATTCACCGCATGTGTGCCTCCTACTCTGAACTCGAGCTT
 GTGACCTCAGCTGAAGGTCTGAACAGCTCTCAAAGCTGGCCTGCCTCATTGGCGTGNAGGG
 TGGTCACTCACTGGACAGCAGCCTCTCTGTGCTGCGCAGTTTCTATGTGCTGGGGGTGCGCT
 ACCTGACACTTACCTTCACCTGCAGTACCCATGGGCAGAGAGTTCCACCAAGTTCAGACAC
 CACATGTACACCAACGTCAGCGGATTGACAAGCTTTGGTGAGAAAGTAGTAGAGGAGTTGAA
 CCGCCTGGGCATGATGATAGATTTGTCCATATGCATCGGACACCTTGATAAGAAGGGTCCTGG
 AAGTGTCTCAGGCTCCTGTGATCTTCTCCCACTCAGCTGCCAGAGCTGTGTGTGACAATTTG
 TTGAATGTTCCCGATGATATCCTGCAGCTTCTGAAGAACGGTGGCATCGTGATGGTGACACT
 GTCCATGGGGGTGCTGCAGTGCAACCTGCTTGCTAACGTGTCCACTGTGGCAGATCACTTTG
 ACCACATCAGGGCAGTCATTGGATCTGAGTTCATCGGGATTGGTGGAAATTATGACGGGACT
 GGCCGGTTCCCTCAGGGGCTGGAGGATGTGTCCACATACCCAGTCCTGATAGAGGAGTTGCT
 GAGTCGTASCTGGAGCGAGGAAGAGCTTCAAGGTGTCCTTCGTGGAAACCTGCTGCGGGTCT
 TCAGACAAGTGGAAGGTGAGAGAGGAGAGCAGGGCGCAGAGCCCCGTGGAGGCTGAGTTT
 CCATATGGGCAACTGAGCACATCCTGCCACTCCCACCTCGTGCCTCAGAATGGACACCAGGC
 TACTCATCTGGAGGTGACCAAGCAGCCAACCAATCGGGTCCCCTGGAGGTCTCAAATGCCT
 CCCCATACCTTGTTCCAGGCCTTGTTGGCTGCTGCCACCATCCCAACCTTCACCCAGTGGCTC
 TGCT**TGA**CACAGTCGGTCCCCGCGAGGGTCACTGTGGCAAAGCCTCACAAAGCCCCCTCTCCT
 AGTTCATTCACAAGCATATGCTGAGAATAAACATGTTACACATGGAAAA

FIGURE 38

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59817

><subunit 1 of 1, 487 aa, 1 stop, 2 unknown

><MW: 53569.32, pI: 7.68, NX(S/T): 5

MQPTGREGSRALSRRYLRRLLLLLLLLLLLLLRQPVTRAETTPGAPRALSTLGSPSLFTTPGVPS
ALTTPGLTTPGTPKTLDLRGRAQALMRSFPLVDGHNDLPQVLRQRYKNVLQDVNLRNFSHGQ
TSLDRLRDGLVGAQFWSASVSCQSQDQTAVRLALEQIDLIHRMCASYSELELV TSAEGLNSS
QKLACLIGVXGGHSLDSSLSVLR SFYVLGVRYLTLTFTCSTPWAESSTKFRHHMYTNVSGLT
SFGEKVVEELNRLGMMIDLSYASDTLIRRVLEVSQAPVIFSHSAARAVCDNLLNVPDDILQL
LKNGGIVMVTLSMGVLQCNLLANVSTVADHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLEDV
STYPVLIIEELLSRXWSEEELQGVLRGNLLRVFRQVEKVVREESRAQSPVEAEFPYGQLSTSCH
SHLVPQNGHQATHLEVTKQPTNRPWRSSNASPYLVPGLVAAATIPTFTQWLC

Important features of the protein:**Signal peptide:**

amino acids 1-36

Transmembrane domain:

amino acids 313-331

N-glycosylation sites.

amino acids 119-122, 184-187, 243-246 and 333-336

N-myristoylation sites.

amino acids 41-46, 59-64, 73-78, 133-138, 182-187, 194-199, 324-329, 354-359, 357-362, 394-399, 427-432 and 472-477.

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 136-146

FIGURE 39

TGCTAGGCTCTGTCCCACAATGCACCCGAGAGCAGGAGCTGAAAGCCTCTAACACCCACAGA
TCCCTCTATGACTGCAATGTGAGGTGTCCGGCTTTGCTGGCCCAGCAAGCCTGATAAGC**ATG**
AAGCTCTTATCTTTGGTGGCTGTGGTCGGGTGTTTGCTGGTGCCCCCAGCTGAAGCCAACAA
GAGTTCTGAAGATATCCGGTGCAAATGCATCTGTCCACCTTATAGAAACATCAGTGGGCACA
TTTACAACCAGAATGTATCCCAGAAGGACTGCAACTGCCTGCACGTGGTGGAGCCCATGCCA
GTGCCTGGCCATGACGTGGAGGCCTACTGCCTGCTGTGCGAGTGCAGGTACGAGGAGCGCAG
CACCACCACCATCAAGGTCATCATTGTCATCTACCTGTCCGTGGTGGGTGCCCTGTTGCTCT
ACATGGCCTTCCTGATGCTGGTGGACCCTCTGATCCGAAAGCCGGATGCATACACTGAGCAA
CTGCACAATGAGGAGGAGAATGAGGATGCTCGCTCTATGGCAGCAGCTGCTGCATCCCTCGG
GGGACCCCGAGCAAACACAGTCCTGGAGCGTGTGGAAGGTGCCCAGCAGCGGTGGAAGCTGC
AGGTGCAGGAGCAGCGGAAGACAGTCTTCGATCGGCACAAGATGCTCAGC**TAGAT**GGGCTGG
TGTGGTTGGGTCAAGGCCCCAACACCATGGCTGCCAGCTTCCAGGCTGGACAAAGCAGGGGG
CTACTTCTCCCTTCCTCGGTTCCAGTCTTCCCTTTAAAGCCTGTGGCATTTCCTCCTT
CTCCCTAACTTTAGAAATGTTGTACTTGGCTATTTTGATTAGGGAAGAGGGATGTGGTCTCT
GATCTCTGTTGTCTTCTTGGGTCTTTGGGGTTGAAGGGAGGGGGAAGGCAGGCCAGAAGGGA
ATGGAGACATTCGAGGCGGCCTCAGGAGTGGATGCGATCTGTCTCTCCTGGCTCCACTCTTG
CCGCCTTCAGCTCTGAGTCTTGGGAATGTTGTTACCCTTGGAAGATAAAGCTGGGTCTTCA
GGAACCTCAGTGTCTGGGAGGAAAGCATGGCCCAGCATTCAGCATGTGTTTCCTTTCTGCAGTG
GTTCTTATCACCACCTCCCTCCCAGCCCCGGCGCCTCAGCCCCAGCCCCAGCTCCAGCCCTG
AGGACAGCTCTGATGGGAGAGCTGGGCCCCCTGAGCCCACTGGGTCTTCAGGGTGCACTGGA
AGCTGGTGTTCGCTGTCCCTGTGCACTTCTCGCACTGGGGCATGGAGTGCCCATGCATACT
CTGCTGCCGGTCCCCTCACCTGCACTTGAGGGGTCTGGGCAGTCCCTCCTCTCCCCAGTGTC
CACAGTCACTGAGCCAGACGGTCGGTTGGAACATGAGACTCGAGGCTGAGCGTGGATCTGAA
CACCACAGCCCCCTGTACTTGGGTGCGCTCTTGTCCCTGAACTTCGTTGTACCAGTGCATGGA
GAGAAAATTTTGTCTCTTGTCTTAGAGTTGTGTGTAATCAAGGAAGCCATCATTAATTG
TTTTATTTCTCTCA

FIGURE 40

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60278

<subunit 1 of 1, 183 aa, 1 stop

<MW: 20574, pI: 6.60, NX(S/T): 3

MKLLSLVAVVGCLLVPPAEANKSSEDIRCKCICPPYRNISGHIYNQNVSQKDCNCLHVVEPM

PVPGHDVEAYCLLCECRYEERSTTTIKVIVIVYLSVVGALLLYMAFLMLVDPLIRKPDAYTE

QLHNEEENEDARSMAAAAASLGGPRANTVLERVEGAQQRWKLQVQEQRKTVFDRHKMLS

Important features:**Signal peptide:**

amino acids 1-20

Transmembrane domain:

amino acids 90-112

N-glycosylation sites.

amino acids 21-24, 38-41 and 47-50

FIGURE 41

AGCGGGTCTCGCTTGGGTTCCGCTAATTTCTGTCTGAGGCGTGAGACTGAGTTCATAGGGTCTGGGTCCCCGA
 ACCAGGAAGGGTTGAGGGAACACAATCTGCAAGCCCCCGCGACCCCAAGTGAGGGGGCCCCGTGTTGGGGTCTCCC
 TCCCTTTGCATTCCCACCCCTCCGGGCTTTGCGTCTTCTGGGGACCCCTCGCCGGGAGATGCGCGCTTGATG
 CGGAGCAAGGATTTCGTCTGCTGCCTGCTCCTACTGGCCGCGGTGCTGATGGTGGAGAGCTCACAGATCGGCAGT
 TCGCGGGCCAACTCAACTCCATCAAGTCCTCTCTGGGCGGGGAGACGCCTGGTCAGGCCGCCAATCGATCTGCG
 GGCATGTACCAAGGACTGGCATTCCGGCGGCAGTAAGAAGGGCAAAAACCTGGGGCAGGCCTACCCTTGTCAGT
 GATAAGGAGTGTGAAGTTGGGAGGTATTGCCACAGTCCCCACCAAGGATCATCGGCCTGCATGGTGTGTCGGAGA
 AAAAGAAGCGCTGCCACCGAGATGGCATGTGCTGCCCGAGTACCCGCTGCAATAATGGCATCTGTATCCCAGTT
 ACTGAAAGCATCTTAACCCCTCACATCCCGCTCTGGATGGTACTCGGCACAGAGATCGAAACCACGGTCATTAC
 TCAAACCATGACTTGGGATGGCAGAATCTAGGAAGACCACACACTAAGATGTCACATATAAAAGGGCATGAAGGA
 GACCCCTGCCTACGATCATCAGACTGCATTGAAGGGTTTTGCTGTGCTCGTCATTTCTGGACCAAAATCTGCAAA
 CCAGTGTCTCCATCAGGGGGAAAGTCTGTACCAACAACGCAAGAAGGGTTCTCATGGGCTGGAAATTTCCAGCGT
 TGGGACTGTGCGAAGGGCCTGTCTTGCAAAGTATGGAAGATGCCACCTACTCCTCCAAAGCCAGACTCCATGTG
 TGTGAGAAATTTGATCACCATTGAGGAACATCATCAATTGCAGACTGTGAAGTTGTGTATTTAATGCATTATAG
 CATGGTGGAAAAATAAGGTTTCAGATGCAGAAGAATGGCTAAAATAAGAAACGTGATAAGAATATAGATGATCACAA
 AAAGGGAGAAAGAAAACATGAAGTGAATAGATTAGAATGGGTGACAAATGCAGTGCAGCCAGTGTTCCTATTATG
 CAACTTGTCTATGTAAATAATGTACACATTTGTGGAAAATGCTATTATTAAGAGAACAAGCACACAGTGGAAAT
 ACTGATGAGTAGCATGTGACTTTCCAAGAGTTTAGGTTGTGCTGGAGGAGAGGTTTCCTTCAGATTGCTGATTGC
 TTATACAAATAACCTACATGCCAGATTTCTATTCAACGTTAGAGTTTAAACAAATACTCCTAGAATAACTGTTA
 TACAATAGGTTCTAAAAATAAAATGCTAAACAAGAAATGAAAACATGGAGCATTGTTAATTTACAACAGAAAAT
 TACCTTTTGATTTGTAACTACTTTCTGCTGTTCAATCAAGAGTCTTGGTAGATAAGAAAAAATCAGTCAATAT
 TTCCAAATAATTGCAAAATAATGGCCAGTTGTTTAGGAAGGCCTTTAGGAAGACAAATAAATAACAAACAAACAG
 CCACAAATACTTTTTTTTCAAATTTTAGTTTTACCTGTAATTAATAAGAACTGATACAAGACAAAAACAGTTCC
 TTCAGATTCTACGGAATGACAGTATATCTCTCTTTATCCTATGTGATTCCTGCTCTGAATGCATTATATTTTCCA
 AACTATACCCATAAATTTGTGACTAGTAAATACTTACACAGAGCAGAATTTTACAGATGGCAAAAAAATTTAAA
 GATGTCCAATATATGTGGGAAAAGAGCTAACAGAGAGATCATTATTTCTTAAAGATTGGCCATAACCTATATTTT
 GATAGAATTAGATTGGTAAATACATGTATTACATACATACTCTGTGGTAATAGAGACTTAAGCTGGATCTGTACTG
 CACTGGAGTAAGCAAGAAAATTTGGGAAAATTTTTCGTTTTGTTTCAGGTTTTGGCAACACATAGATCATATGTCTG
 AGGCACAAGTTGGCTGTTTCATCTTTGAAACCAGGGGATGCACAGTCTAAATGAATATCTGCATGGGATTTGCTAT
 CATAATATTTACTATGCAGATGAATTCAGTGTGAGGTCTGTGTCCGTACTATCCTCAAATTATTTATTTTATAG
 TGCTGAGATCCTCAAATAATCTCAATTTAGGAGGTTTACAAAATGTACTCCTGAAGTAGACAGAGTAGTGAGG
 TTTTATTGCCAAAGGGCTAGTTTTCGGTTTTCTGCAGCCATTGCGGTTAAAAAATATAAGTAGGATAACTTGTA
 ACCTGCATATTGCTAATCTATAGACACCACAGTTTCTAAATTTCTTTGAAACCACTTTACTACTTTTTTTAACTT
 AACTCAGTTCTAAATACTTTGTCTGGAGCACAAAAACAATAAAAGGTTATCTTATAGTCGTGACTTTAACTTTTG
 TAGACCACAATTCACCTTTTGTGTTTTCTTTACTTAAATCCCATCTGCAGTCTCAAATTTAAGTTCTCCAGTAG
 AGATTGAGTTTGAGCCTGTATATCTATTAATAAATTTCAACTTCCACATATATTTACTAAGATGATTAAGACTTA
 CATTTTCTGCACAGGTCTGCAAAAAACAAAAATTATAAACTAGTCCATCCAAGAACCAGTTTGTATAAACAGGT
 TGCTATAAGCTTGTGAAATGAAAATGGAACATTTCAATCAAACATTTCTATATAACAATTATTATATTTACAAT
 TTGGTTTTCTGCAATATTTTTCTTATGTCCACCCTTTAAAAATTATTATTTGAAGTAATTTATTTACAGGAATG
 TTAATGAGATGTATTTTCTTATAGAGATATTTCTTACAGAAAGCTTTGTAGCAGAATATATTTGCAGCTATTGAC
 TTTGTAATTTAGGAAAAATGTATAATAAGATAAAATCTATTAAATTTTTCTCTCTAAAAACTGAAAAA
 AAAAAAAAAAAAAAAAAA

FIGURE 42

MAALMRSKDSSCCLLLLLAAVLMVESSQIGSSRAKLNSIKSSLGGETPGQAANRSAGMYQGLA
FGGSKKGKNLGQAYPCSSDKECEVGRYCHSPHQSSACMVCRRKKKRCHRDGMCCPSTRCNN
GICIPVTESILTPHIPALDGTRHRDRNHGHYSNHDLGWQNLGRPHTKMSHIKGHEGDPCLRS
SDCIEGFCCARHFWTKICKPVLHQGEVCTKQRKKGSHGLEIFQRCDCAKGLSCKVWKDATYS
SKARLHVCQKI

Signal peptide:

amino acids 1-25

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 43

TCTCAATCTGCTGACCTCGTGATCCGCCTGACCTTGTAATCCACCTACCTTGGCCTCCCAA
 GTGTTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCAACATCACGTTTTTAAAAATTGATT
 TCTTCAAATTCATGGCAAATATTTCCCTTCCCTTTAACTTCTTATGTCAGAATGAGGAAGGA
 TAGCTGCATTTATTTAGTCAGTTTTTCATTGCATAGTAATATTTTCATGTAGTATTTTCTAAG
 TTATATTTTAGTAATTCATATGTTTTAGATTATAGGTTTTAACATACTTGTGAAAATACTTG
ATGTGTTTTAAAGCCTTGGGCAGAAATTCTGTATTGTTGAGGATTTGTTCTTTTATCCCCCT
 TTAAAGTCATCCGTCCTTGGCTCAGGATTTGGAGAGCTTGCACCACCAAAAATGGCAAACA
 TCACCAGCTCCCAGATTTTGGACCAGTTGAAAGCTCCGAGTTTGGGCCAGTTTACCACCACC
 CCAAGTACACAGCAGAATAGTACAAGTCACCCTACAACCTACTACTTCTTGGGACCTCAAGCC
 CCCAACATCCCAGTCCTCAGTCCTCAGTCATCTTGACTTCAAATCTCAACCTGAGCCATCCC
 CAGTTCTTAGCCAGTTGAGCCAGCGACAACAGCACCAGAGCCAGGCAGTCACTGTTCCCTCCT
 CCTGGTTTGGAGTCCTTTCCTTCCCAGGCAAACTTCGAGAATCAACACCTGGAGACAGTCC
 CTCCACTGTGAACAAGCTTTTGCAGCTTCCCAGCACGACCATTGAAAATATCTCTGTGTCTG
 TCCACCAGCCACAGCCCAAACACATCAAACCTTGCTAAGCGGCGGATACCCCCAGCTTCTAAG
 ATCCCAGCTTCTGCAGTGGAAATGCCTGGTTCAGCAGATGTCACAGGATTAATGTGCAGTT
 TGGGGCTCTGGAATTTGGGTCAGAACCTTCTCTCTCTGAATTTGGATCAGCTCCAAGCAGTG
 AAAATAGTAATCAGATTCCCATCAGCTTGTATTCTGAAGTCTTTAAGTGAGCCTTTGAATACA
 TCTTTATCAATGACCAGTGCAGTACAGAACTCCACATATACAACCTCCGTCATTACCTCCTG
 CAGTCTGACAAGCTCATCACTGAATTCTGCTAGTCCAGTAGCAATGTCTTCCTCTTATGACC
 AGAGTTCTGTGCATAACAGGATCCCATACCAAAGCCCTGTGAGTTCATCAGAGTCAGCTCCA
 GGAACCATCATGAATGGACATGGTGGTGGTTCGAAGTCAGCAGACACTAGACAGTAAGTATAG
 CAGCAAGCTACTCTTGTTCATGGCTGGTGCCAACCAAACAGAGGAAGAGGATAGCTCACGTGA
 TGTGGAAAACACCAGTTGGTCAATGGCTCATTCGTTAAAAAAGCAGCCCTTTTGCTTTTTTGT
 TTTTGGACCAGGTGTTGGCTGTGGTGTATTAGAAATGTCTTAACCACAGCAAGAAGGAGGT
 GGTGGTCTCATATTCTTCTGCCCTAATCAGACTGCACCACAAGTGCAGCATAACAGTATGCAT
 TTAAAGATGCTTGGGCCAGGCGGGGTGGCTGATGCCATAATCCCAGTGCTTTGGGGGGCC
 AAGGCAGGCAGATTGCCCAAGCTCAGGAGTTTGAAGACCACCTGGGCAACATGGTGAACTC
 TGTCTCTACTAAAATACGAAAACTAGCCGGGTGTGGTGGCGGCGCGTGCCTGTAATCCCAG
 CTACTTGGGAGGCTGAGGCACAAGAATCGCTTGAGCCAGCTTGGGCTACAAAGTGAGACTCC
 GTCTGAAAAGA

FIGURE 44

MCFKALGRNSVLLRICSFIPLLKSSVLGSGFGELAPPKMANITSSQILDQLKAPSLGQFTTT
PSTQQNSTSHPTTTTTSWDLKPPTSQSSVLSHLDFKSQPEPSPVLSQLSQRQQHQSQAVTVPP
PGLESFPSQAKLRESTPGDSPSTVNKLLQLPSTTIENISVSVHQPPKHIKLAKRRIPPASK
IPASAVEMPGSADVTGLNVQFGALEFGSEPSLSEFGSAPSSSENSNQIPISLYSKSLSEPLNT
SLSMTSAVQNSTYTTTSVITSCSLTSSSLNSASPVAMSSSYDQSSVHNRIPTYQSPVSSSESAP
GTIMNGHGGGRSQQTLDISKYSSKLLLSWLVP TKQRKRIAHVMWKTPVGQWLIR

Signal peptide:

amino acids 1-24

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

FIGURE 45

GCCGAGTGGGACAAAGCCTGGGGCTGGGCGGGGGCC**ATG**GCGCTGCCATCCCGAATCCTGCT
 TTGGAAACTTGTGCTTCTGCAGAGCTCTGCTGTTCTCCTGCACTCAGCGGTGGAGGAGACGG
 ACGCGGGGCTGTACACCTGCAACCTGCACCATCACTACTGCCACCTCTACGAGAGCCTGGCC
 GTCCGCCTGGAGGTCACCGACGGCCCCCGGCCACCCCGCCTACTGGGACGGCGAGAAGGA
 GGTGCTGGCGGTGGCGCGCGGGCGCACCCGCGCTTCTGACCTGCGTGAACCGCGGGCACGTGT
 GGACCGACCGGCACGTGGAGGAGGCTCAACAGGTGGTGCCTGGGACCGGCAGCCGCCCGGG
 GTCCCGCACGACCGCGCGGACCGCCTGCTGGACCTCTACGCGTCGGGCGAGCGCCGCGCCTA
 CGGGCCCCCTTTTTCTGCGCGACCGCGTGGCTGTGGGCGCGGATGCCTTTGAGCGCGGTGACT
 TCTCACTGCGTATCGAGCCGCTGGAGGTCGCCGACGAGGGCACCTACTCCTGCCACCTGCAC
 CACCATTA CTGTGGCCTGCACGAACGCCGCGTCTTCCACCTGACGGTCGCCGAACCCACGC
 GGAGCCGCCCCCGGGGCTCTCCGGGCAACGGCTCCAGCCACAGCGGCGCCCCAGGCCACG
 ACCCCACACTGGCGCGCGGCCACAACGTCATCAATGTCATCGTCCCCGAGAGCCGAGCCCAC
 TTCTTCCAGCAGCTGGGCTACGTGCTGGCCACGCTGCTGCTCTTCATCCTGCTACTGGTCAC
 TGTCTCCTGGCCGCCCGCAGGCGCCGCGGAGGCTACGAATACTCGGACCAGAAGTCGGGAA
 AGTCAAAGGGGAAGGATGTTAACTTGGCGGAGTTGCTGTGGCTGCAGGGGACCAGATGCTT
 TACAGGAGTGAGGACATCCAGCTAGATTACAAAAACAACATCCTGAAGGAGAGGGCGGAGCT
 GGCCACAGCCCCCTGCCTGCCAAGTACATCGACCTAGACAAAGGGTTCCGGAAGGAGAACT
 GCAAA**TAG**GGAGGCCCTGGGCTCCTGGCTGGGCCAGCAGCTGCACCTCTCCTGTCTGTGCTC
 CTCGGGGCATCTCCTGATGCTCCGGGGCTCACCCCCCTTCCAGCGGCTGGTCCCGCTTTCTCCT
 GGAATTTGGCCTGGGCGTATGCAGAGGCCGCTCCACACCCCTCCCCAGGGGCTTGGTGGC
 AGCATAGCCCCACCCCTGCGGCCTTTGCTCACGGGTGGCCCTGCCACCCCTGGCACAAACC
 AAAATCCCACTGATGCCATCATGCCCTCAGACCCCTTCTGGGCTCTGCCCCGCTGGGGGCTG
 AAGACATTCCTGGAGGACACTCCCATCAGAACCTGGCAGCCCCAAAACCTGGGGTCAGCCTCA
 GGGCAGGAGTCCCACTCCTCCAGGGCTCTGCTCGTCCGGGGCTGGGAGATGTTCTGGAGGA
 GGACACTCCCATCAGAACTTGGCAGCCTTGAAGTTGGGGTCAGCCTCGGCAGGAGTCCCACT
 CCTCCTGGGGTGCTGCCTGCCACCAAGAGCTCCCCACCTGTACCACCATGTGGGACTCCAG
 GCACCATCTGTTCTCCCCAGGGACCTGCTGACTTGAATGCCAGCCCTTGCTCCTCTGTGTTG
 CTTTGGGCCACCTGGGGCTGCACCCCCCTGCCCTTTCTCTGCCCCATCCCTACCCTAGCCTTG
 CTCTCAGCCACCTTGATAGTCACTGGGCTCCCTGTGACTTCTGACCCTGACACCCCTCCCTT
 GGA CTCTGCCTGGGCTGGAGTCTAGGGCTGGGGCTACATTTGGCTTCTGTACTGGCTGAGGA
 CAGGGGAGGGAGTGAAGTTGGTTTGGGGTGGCCTGTGTTGCCACTCTCAGCACCCACATTT
 GCATCTGCTGGTGGACCTGCCACCATCACAATAAAGTCCCCATCTGATTTTTAAAAA
 AAAAA

FIGURE 46

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60618

<subunit 1 of 1, 341 aa, 1 stop

<MW: 38070, pI: 6.88, NX(S/T): 1

MALPSRILLWKLVLQLQSSAVLLHSAVEETDAGLYTCNLHHHYCHLYESLAVRLEVTDGPPAT
 PAYWDGEKEVLAVARGAPALLTCVNRGHVWTD RHVEEAQQV VHWDRQPPGVPHDRADRLLDL
 YASGERRAYGPLFLRDRVAVGADAFERGFSLRIEPLVAD EGTYSCHLHHHYCGLHERRVF
 HLTVAEPHAEP PPRGSPGNGSSSHSGAPGPDPTLARGHNVINVIVPESRAHFFQQLGYVLATL
 LLFILLLVTVLLAARRRRGGYEYSDQKSGKSKGKDVNLAEFAVAAGDQMLYRSEDIQLDYKN
 NILKERAELAHSPLPAKYIDL DKGFRKENCK

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 237-262

N-glycosylation site.

amino acids 205-208

Cell attachment sequence.

amino acids 151-154

Coproporphyrinogen III oxidase proteins.

amino acids 115-140

FIGURE 47

CGCCGGAGGCAGCGGCGGCGTGGCGCAGCGGCGAC**ATG**GCCGTTGTCTCAGAGGACGACTTT
CAGCACAGTTCAAACCTCACCTACGGAACCACAAGCAGCAGTCTCCGAGCTGACCAGGAGGC
ACTGCTTGAGAAGCTGCTGGACCGCCCGCCCCCTGGCCTGCAGAGGCCCGAGGACCGCTTCT
GTGGCACATACATCATCTTCTTCAGCCTGGGCATTGGCAGTCTACTGCCATGGAACTTCTTT
ATCACTGCCAAGGAGTACTGGATGTTCAAACCTCCGCAACTCCTCCAGCCCAGCCACCGGGGA
GGACCCTGAGGGCTCAGACATCCTGAACTACTTTGAGAGCTACCTTGCCGTTGCCTCCACCG
TGCCCTCCATGCTGTGCCTGGTGGCCAACTTCCTGCTTGTCAACAGGGTTGCAGTCCACATC
CGTGTCTGGCCTCACTGACGGTCATCCTGGCCATCTTCATGGTGATAACTGCACTGGTGAA
GGTGGACACTTCCTCCTGGACCCGTGGTTTTTTTTGCGGTCACCATTGTCTGCATGGTGATCC
TCAGCGGTGCCTCCACTGTCTTCAGCAGCAGCATCTACGGCATGACCGGCTCCTTTCTATG
AGGAACTCCCAAGCACTGATATCAGGAGGAGCCATGGGCGGGACGGTCAGCGCCGTGGCCTC
ATTGGTGGACTTGGCTGCATCCAGTGATGTGAGGAACAGCGCCCTGGCCTTCTTCCTGACGG
CCACCATCTTCCTCGTGCTCTGCATGGGACTCTACCTGCTGCTGTCCAGGCTGGAGTATGCC
AGGTACTACATGAGGCCTGTTCTTGCGGCCCATGTGTTTTCTGGTGAAGAGGAGCTTCCCCA
GGACTCCCTCAGTGCCCCCTTCGGTGGCCTCCAGATTCATTGATTCCCACACACCCCCCTCTCC
GCCCCATCCTGAAGAAGACGGCCAGCCTGGGCTTCTGTGTACCTACGTCTTCTTCATCACC
AGCCTCATCTACCCCGCCGTCTGCACCAACATCGAGTCCCTCAACAAGGGCTCGGGCTCACT
GTGGACCACCAAGTTTTTTCATCCCCCTCACTACCTTCCTCCTGTACAACCTTTGCTGACCTAT
GTGGCCGGCAGCTCACCGCCTGGATCCAGGTGCCAGGGCCCAACAGCAAGGCGTCCCAGGG
TTCGTGCTCCTCCGGACCTGCCTCATCCCCCTCTTCGTGCTCTGTAACCTACCAGCCCCGCGT
CCACCTGAAGACTGTGGTCTTCCAGTCCGATGTGTACCCCGCACTCCTCAGCTCCCTGCTGG
GGCTCAGCAACGGCTACCTCAGCACCTGGCCCTCCTCTACGGGCCTAAGATTGTGCCCAGG
GAGCTGGCTGAGGCCACGGGAGTGGTGATGTCCTTTTATGTGTGCTTGGGCTTAACACTGGG
CTCAGCCTGCTCTACCCTCCTGGTGCACCTCATC**TAGA**AGGGGAGGACACAAGGACATTGGTG
CTTCAGAGCCTTTGAAGATGAGAAGAGAGTGCAGGAGGGCTGGGGGCCATGGAGGAAAGGCC
TAAAGTTTCACTTGGGGACAGAGAGCAGAGCACACTCGGGCCTCATCCCTCCCAAGATGCCA
GTGAGCCACGTCCATGCCCATTCCGTGCAAGGCAGATATTCCAGTCATATTAACAGAACT
CCTGAGACAGTTGAAGAAGAAATAGCACAAATCAGGGGTACTCCCTTCACAGCTGATGGTTA
ACATTCCACCTTCTTTCTAGCCCTTCAAAGATGCTGCCAGTGTTGCCCCTAGAGTTATTACA
AAGCCAGTGCCAAAACCCAGCCATGGGCTCTTTGCAACCTCCCAGCTGCGCTCATTCCAGCT
GACAGCGAGATGCAAGCAAATGCTCAGCTCTCCTTACCCTGAAGGGGTCTCCCTGGAATGGA
AGTCCCCTGGCATGGTCAGTCTCAGGCCCAAGACTCAAGTGTGCACAGACCCCTGTGTTCT
GCGGGTGAACAACTGCCCACTAACCAGACTGGAAAACCCAGAAAGATGGGCCTTCCATGAAT
GCTTCATTCCAGAGGGACAGAGGGCCTCCCTGTGCAAGGGATCAAGCATGTCTGGCCTGGG
TTTTCAAAAAAAGAGGGATCCTCATGACCTGGTGGTCTATGGCCTGGGTCAAGATGAGGGTC
TTTCAGTGTTCTGTTTACAACATGTCAAAGCCATTGGTTCAAGGGCGTAATAAATACTTGC
GTATTCAAAAA

FIGURE 48

MAVVSEDDFQHSSNSTYGTSSSLRADQEALLEKLLDRPPPGQLQRPEDRFCGTYYIFFSLGI
GSLLPWNFFITAKEYWMFKLRNSSSPATGEDPEGSDILNYFESYLAVASTVPSMLCLVANFL
LVNRVAVHIRVLASLTVILAI FMVITALVKVDTSSWTRGFFAVTIVCMVILSGASTVFSSSI
YGMTGSFPMRNSQALISGGAMGGTVSAVASLVDLAASSDVRNSALAFFLTATIFLVLCMGly
LLLSRLEYARYYMRPVLA AHVFSGEEELPQDSL SAPSVASRFIDSHTPPLRPILKKTASLGF
CVTYVFFITS LIYPAVCTNIESLNKSGSLWTTKFFIPLTTFLLYNFADLCGRQLTAWIQVP
GPNSKALPGFVLLRTCLIPLFVLCNYQPRVHLKTVVFQSDVYPALLSSLLGLSNGYLSTLAL
LYGPKIVPRELAEATGVVMSFYVCLGLTLGSACSTLLVHLI

Transmembrane domain:

amino acids 50-74 (type II), 105-127, 135-153, 163-183, 228-252,
305-330, 448-472

FIGURE 49

GACAGTGGAGGGCAGTGGAGAGGACCGCGCTGTCTTGCTGTCACCAAGAGCTGGAGACACCA
 TCTCCCACCGAGAGTCA**ATG**GCCCCATTGGCCCTGCACCTCCTCGTCCTCGTCCCCATCCTCC
 TCAGCCTGGTGGCCTCCCAGGACTGGAAGGCTGAACGCAGCCAAGACCCCTTCGAGAAATGC
 ATGCAGGATCCTGACTATGAGCAGCTGCTCAAGGTGGTGAACCTGGGGGCTCAATCGGACCCT
 GAAGCCCCAGAGGGTGATTGTGGTTGGCGCTGGTGTGGCCGGGCTGGTGGCCGCCAAGGTGC
 TCAGCGATGCTGGACACAAGGTACCATCCTGGAGGCAGATAACAGGATCGGGGGCCGCATC
 TTCACCTACCGGGACCAGAACACGGGCTGGATTGGGGAGCTGGGAGCCATGCGCATGCCCAG
 CTCTCACAGGATCCTCCACAAGCTCTGCCAGGGCCTGGGGCTCAACCTGACCAAGTTCACCC
 AGTACGACAAGAACACGTGGACGGAGGTGCACGAAGTGAAGCTGCGCAACTATGTGGTGGAG
 AAGGTGCCCCGAGAAGCTGGGCTACGCCTTGCGTCCCCAGGAAAAGGGCCACTCGCCCCGAAGA
 CATCTACCAGATGGCTCTCAACCAGGCCCTCAAAGACCTCAAGGCACTGGGCTGCAGAAAGG
 CGATGAAGAAAGTTTGAAAGGCACACGCTCTTGGAATATCTTCTCGGGGAGGGGAACCTGAGC
 CGGCCGGCCGTGCAGCTTCTGGGAGACGTGATGTCCGAGGATGGCTTCTTCTATCTCAGCTT
 CGCCGAGGCCCTCCGGGGCCACAGCTGCCTCAGCGACAGACTCCAGTACAGCCGCATCGTGG
 GTGGCTGGGACCTGCTGCCGCGCGCGCTGCTGAGCTCGCTGTCCGGGCTTGTGCTGTTGAAC
 GCGCCCGTGGTGGCGATGACCCAGGGACCGCACGATGTGCACGTGCAGATCGAGACCTCTCC
 CCCGGCGCGGAATCTGAAGGTGCTGAAGGCCGACGTGGTGCTGCTGACGGCGAGCGGACCGG
 CGGTGAAGCGCATCACCTTCTCGCCGCCGTGCCCGCCACATGCAGGAGGCGCTGCGGAGG
 CTGCACTACGTGCCGGCCACCAAGGTGTTCCCTAAGCTTCCGCAGGCCCTTCTGGCGCGAGGA
 GCACATTGAAGGCGGCCACTCAAACACCGATCGCCCGTCGCGCATGATTTTCTACCCGCCGC
 CGCGCGAGGGCGCGCTGCTGCTGGCCTCGTACACGTGGTTCGGACGCGGCGGCAGCGTTCCGCC
 GGCTTGAGCCGGGAAGAGGCGTTGCGCTTGGCGCTCGACGACGTGGCGGCATTGCACGGGCC
 TGTCGTGCGCCAGCTCTGGGACGGCACCGGCGTCGTCAAGCGTTGGGCGGAGGACCAGCACA
 GCCAGGGTGGCTTTGTGGTACAGCCGCCGGCGCTCTGGCAAACCGAAAAGGATGACTGGACG
 GTCCCTTATGGCCGCATCTACTTTGCCGGCGAGCACACCGCCTACCCGCACGGCTGGGTGGA
 GACGGCGGTCAAGTCGGCGCTGCGCGCCGCCATCAAGATCAACAGCCGGAAGGGGCCCTGCAT
 CGGACACGGCCAGCCCCGAGGGGCACGCATCTGACATGGAGGGGCAGGGGCATGTGCATGGG
 GTGGCCAGCAGCCCCTCGCATGACCTGGCAAAGGAAGAAGGCAGCCACCCTCCAGTCCAAGG
 CCAGTTATCTCTCCAAAACACGACCCACACGAGGACCTCGCAT**TAA**AGTATTTTCGGAAAAA
 AA

FIGURE 50

MAPLALHLLVLPILLSLVLASQDWKAERSQDPFEKCMQDPDYEQLLKVVWGLNRTLKPQRV
IVVGAGVAGLVAAKVLSDAGHKVTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRI
HKLCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGALRPQEKGHSPEDIYQMA
LNQALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSDGGFFYLSFAEALR
AHSCLSDRLQYSRIVGGWDLPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPPARNL
KVLKADVLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFRRPFWRREEHIEGG
HSNTDRPSRMIFYPPPREGALLASYTWSDAAAAFAGLSREEALRLALDDVAALHGPVVRQL
WDGTGVVKRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWVETAVKS
ALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQ
NTTHTRTSH

Signal peptide:

amino acids 1-21

FIGURE 51

CTGACATGGCCTGACTCGGGACAGCTCAGAGCAGGGCAGAACTGGGGACACTCTGGGCCGGCCTTCTGCCTGCAT
 GAGCGCTCTGAAGCCACCCTGTCTCTGGAGGAACACGAGCGAGGGAAGAAGGACAGGGACTCGTGTGGCAGGAA
 GAACTCAGAGCCGGGAAGCCCCCATTACTAGAAGCACTGAGAGATGCGGCCCCCTCGCAGGGTCTGAATTTCTCT
 GCTGCTGTTACAAAGATGCTTTTTATCTTTAACTTTTTGTTTTCCCCACTTCCGACCCCGGCGTTGATCTGCAT
 CCTGACATTTGGAGCTGCCATCTTCTTGTGGCTGATCACCAGACCTCAACCCGTCTTACCTCTTCTTGACCTGAA
 CAATCAGTCTGTGGGAATTGAGGGAGGAGCACGGAAGGGGGTTCACAGAAAGAACAATGACCTAACAAGTTGCTG
 CTTCTCAGATGCCAAGACTATGTATGAGGTTTTCCAAAGAGGACTCGCTGTGTCTGACAATGGGCCCTGCTTGGG
 ATATAGAAAACCAACCAGCCCTACAGATGGCTATCTTACAAACAGGTGTCTGATAGAGCAGAGTACCTGGGTTT
 CTGTCTCTTGCAATAAAGGTTATAAATCATCACCAGACCAGTTTGTGCGCATCTTTGCTCAGAATAGGCCAGAGTG
 GATCATCTCCGAATTGGCTTGTACACGTACTCTATGGTAGCTGTACCTCTGTATGACACCTTGGGACCAGAAGC
 CATCGTACATATTGTCAACAAGGCTGATATCGCCATGGTGATCTGTGACACACCCCCAAAAGGCATTGGTGCTGAT
 AGGGAATGTAGAGAAAGGCTTCACCCGAGCCTGAAGGTGATCATCCTTATGGACCCCTTTGATGATGACCTGAA
 GCAAAGAGGGGAGAAGAGTGGAATTGAGATCTTATCCCTATATGATGCTGAGAACCTAGGCAAAGAGCACTTCAG
 AAAACCTGTGCCTCCTAGCCCAGAAGACCTGAGCGTCATCTGCTTACCAGTGGGACCACAGGTGACCCCAAAGG
 AGCCATGATAACCCATCAAAATATTGTTTTCAAATGCTGCTGCCTTTCTCAAATGTGTGGAGCATGCTTATGAGCC
 CACTCCTGATGATGTGGCCATATCCTACCTCCCTCTGGCTCATATGTTTGAGAGGATTGTACAGGCTGTTGTGTA
 CAGCTGTGGAGCCAGAGTTGGATTCTTCCAAGGGGATATTCGGTTGCTGGCTGACGACATGAAGACTTTGAAGCC
 CACATTGTTTTCCCGCGGTGCCTCGACTCCTTAACAGGATCTACGATAAGGTACAAAATGAGGCCAAGACACCCTT
 GAAGAAGTTCTTGTGAAGCTGGCTGTTTCCAGTAAATTCAAAGAGCTTCAAAGGGTATCATCAGGCATGATAG
 TTTCTGGGACAAGCTCATCTTTGCAAAGATCCAGGACAGCCTGGGCGGAAGGGTTCGTGTAATTGTCACTGGAGC
 TGCCCCCATGTCCACTTCAGTCATGACATTCTTCCGGGCAGCAATGGGATGTGAGGTGTATGAAGCTTATGGTCA
 AACAGAATGCACAGGTGGCTGTACATTTACATTACCTGGGACTGGACATCAGGTACGTTGGGGTGCCCCCTGGC
 TTGCAATTACGTGAAGCTGGAAGATGTGGCTGACATGAAGTACTTTACAGTGAATAATGAAGGAGAGGTCTGCAT
 CAAGGGTACAAACGTGTTCAAAGGATACCTGAAGGACCCTGAGAAGACACAGGAAGCCCTGGACAGTGATGGCTG
 GCTTCACACAGGAGACATTGGTTCGCTGGCTCCCGAATGGAACCTCTGAAGATCATCGACCGTAAAAAGAACATTTT
 CAAGCTGGCCCAAGGAGAATACATTGCACCAGAGAAGATAGAAAATATCTACAACAGGAGTCAACCAGTGTACAA
 AATTTTTGTACACGGGGAGAGCTTACGGTCATCCTTAGTAGGAGTGGTGGTTCTGACACAGATGTACTTCCCTC
 ATTTGCAGCCAAGCTTGGGGTGAAGGGCTCCTTTGAGGAACTGTGCCAAAACCAAGTTGTAAGGGAAGCCATTTT
 AGAAGACTTGCAGAAAATTGGGAAAGAAAGTGGCCTTAAACTTTTGAACAGGTCAAAGCCATTTTTCTTCATCC
 AGAGCCATTTTCCATTGAAAATGGGCTCTTGACACCAACATTGAAAGCAAAGCGAGGAGAGCTTTCCAAATACTT
 TCGGACCCAAATTGACAGCCTGTATGAGCACATCCAGGATTAGGATAAGGTACTTAAGTACCTGCCGGCCCACTG
 TGCATGCTTGTGAGAAAATGGATTAAAAACTATTCTTACATTTGTTTTGCCTTTTCTCCTATTTTTTTTTAACC
 TGTTAAACTCTAAAGCCATAGCTTTTTGTTTTATATTGAGACATATAATGTGTAAACTTAGTTCCCAAATAAATCA
 ATCCTGTCTTTCCCATCTTCGATGTTGCTAATATTAAGGCTTCAGGGCTACTTTTATCAACATGCCTGTCTTCAA
 GATCCCAGTTTATGTTCTGTCTCCTTCCCTCATGATTTCCAACCTTAATACTATTAGTAACCACAAGTTCAAGGGT
 CAAAGGGACCCCTCTGTGCCTTCTTCTTTGTTTTGTGATAAACATAAATTGCCAACAGTCTCTATGCTTATTTACA
 TCTTCTACTGTTCAAATAAGAGATTTTTAAATTCTGAAAACTGCTTACAATTCTGTTTTCTAGCCACTCCAC
 AAACCACTAAAAATTTAGTTTTAGCCTATCACTCATGTCAATCATATCTATGAGACAAATGTCTCCGATGCTCTT
 CTGCGTAAATTAATTTGTGTACTGAAGGGAAAAGTTTGTATCATACCAAACATTTCTTAACTCTCTAGTTAGATA
 TCTGACTTGGGAGTATTAATAATTGGGTCTATGACATACTGTCCAAAAGGAATGCTGTTCTTAAAGCATTATTTA
 CAGTAGGAAGTGGGGAGTAAATCTGTTCCCTACAGTTTGTGCTGAGCTGGAAGCTGTGGGGGAAGGAGTTGACA
 GGTGGGCCCAGTGAACTTTTCCAGTAAATGAAGCAAGCACTGAATAAAAACCTCCTGAACCTGGGAACAAAGATCT
 ACAGGCAAGCAAGATGCCACACAACAGGCTTATTTCTGTGAAGGAACCAACTGATCTCCCCACCCTTGGATT
 AGAGTTCCTGCTCTACCTTACCCACAGATAACACATGTTGTTTTCTACTTGTAAATGTAAAGTCTTTAAAAATAAAC
 TATTACAGATAAAAAA

FIGURE 52

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60775

<subunit 1 of 1, 739 aa, 1 stop

<MW: 82263, pI: 7.55, NX(S/T): 3

MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNLLLLFTKMLFIEN
FLFSPLPTPALICILTFGAAILWLITRPQPVLPDLLNNSVGIEGGARKGVSQKNNDLTS
CCFSDAKTMYEVFQRLAVSDNGPCLGYRKPNQPYRWLSYKQVSDRAEYLGSCLLHKGYKSS
PDQFVGIFAQNRPEWIISELACYTYSMAVAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKAL
VLIGNVEKGFTPSLKVIIILMDPFDDDLKQRGEKSGIEILSLYDAENLGKEHFRKPVPPSPED
LSVICFTSGTTGDPKGAMITHQNIVSNAAAFLKCVEHAYEPTPDDVAISYLPLAHMFERIVQ
AVVYSCGARVGGFFQGDIRLLADDMKTLKPTLFPVPRLLNRIYDKVQNEAKTPLKKFLLKLA
VSSKFKELOQGIIRHDSFWDKLIFAKIQDSLGGVRVIVTGAAPMSTSVMTFFRAAMGCQVY
EAYGQTECTGGCTFTLPGDWTSGHVGVPACNYVKLEDVADMNYFTVNNEGEVCIKGTNVFK
GYLKDPEKTQEALDSDGWLHTGDIGRWLPNGTLKIIDRKKNIKLAQGEYIAPEKIENIYNR
SQPVLQIFVHGESLRSSLVGVVVPDTPDLPSFAAKLGVKGSFEELCQNQVVRVREAILEDLQKI
GKESGLKTFEQVKAI FLHPEPFSEIENGLLTPTLKAKRGELSKYFRTQIDSLYEHIQD

Important features:**Type II transmembrane domain:**

amino acids 61-80

Putative AMP-binding domain signature.

amino acids 314-325

N-glycosylation site.

amino acids 102-105, 588-591 and 619-622

FIGURE 53

GGAGGCGGAGGCGCGGCGAGCCGGGCGAGCAGTGAGGGCCCTAGCGGGGCCCCGAGCGGGG
 CCCGGGGCCCCCTAAGCCATTCTGAAGTCATGGGCTGGCCAGGACATTGGTGACCCGCCAAT
 CCGGT**ATG**GACGACTGGAAGCCCAGCCCCCTCATCAAGCCCTTTGGGGCTCGGAAGAAGCGG
 AGCTGGTACCTTACCTGGAAGTATAAACTGACAAACCAGCGGGCCCTGCGGAGATTCTGTCA
 GACAGGGGCGGTGCTTTTCCTGCTGGTGAAGTGTCAATATCAAGTTGATCCTGGACA
 CTCGGCGAGCCATCAGTGAAGCCAATGAAGACCCAGAGCCAGAGCAAGACTATGATGAGGCC
 CTAGGCCGCTGGAGCCCCACGGCGCAGAGGCAGTGGTCCCCGGCGGGTCCTGGACGTAGA
 GGTGTATTCAAGTCGCAGCAAAGTATATGTGGCAGTGGATGGCACCACGGTGCTGGAGGATG
 AGGCCCGGGAGCAGGGCGGGGCATCCATGTCATTGTCTCAACCAGGCCACGGGCCACGTG
 ATGGCAAACGTGTGTTTGACACGTACTCACCTCATGAGGATGAGGCCATGGTGCTATTCT
 CAACATGGTAGCGCCCGGCGAGTGCTCATCTGCACTGTCAAGGATGAGGGCTCCTTCCACC
 TCAAGGACACAGCCAAGGCTCTGCTGAGGAGCCTGGGCAGCCAGGCTGGCCCTGCCCTGGGC
 TGGAGGGACACATGGGCCTTCGTGGGACGAAAAGGAGGTCTGTCTTCGGGGAGAAACATTC
 TAAGTCACCTGCCCTCTCTTCCTGGGGGGACCCAGTCTGCTGAAGACAGATGTGCCATTGA
 GCTCAGCAGAAGAGGCAGAGTGCCACTGGGCAGACACAGAGCTGAACCGTCGCCGCCGGCGC
 TTCTGCAGCAAAGTTGAGGGCTATGGAAGTGTATGCAGCTGCAAGGACCCACACCCATCGA
 GTTCAGCCCTGACCCACTCCCAGACAACAAGGTCTCAATGTGCCTGTGGCTGTCAATTGCAG
 GGAACCGACCCAATTACCTGTACAGGATGCTGCGCTCTCTGCTTTCAGCCAGGGGGTGTCT
 CCTCAGATGATAACAGTTTTCATTGACGGCTACTATGAGGAACCCATGGATGTGGTGGCACT
 GTTTGGTCTGAGGGGCATCCAGCATACTCCCATCAGCATCAAGAATGCCCGCGTGTCTCAGC
 ACTACAAGGCCAGCCTCACTGCCACTTTCAACCTGTTTCCGGAGGCCAAGTTTGCTGTGGTT
 CTGGAAGAGGACCTGGACATTGCTGTGGATTTTTTCAGTTTCCTGAGCCAATCCATCCACCT
 ACTGGAGGAGGATGACAGCCTGTACTGCATCTCTGCCTGGAATGACCAGGGGTATGAACACA
 CGGCTGAGGACCCAGCACTACTGTACCGTGTGGAGACCATGCCTGGGCTGGGCTGGGTGCTC
 AGGAGGTCTTGTACAAGGAGGAGCTTGAGCCCAAGTGGCCTACACCGGAAAAGCTCTGGGA
 TTGGGACATGTGGATGCGGATGCCTGAACAACGCCGGGGCCGAGAGTGCATCATCCCTGACG
 TTTCCCGATCCTACCACTTTGGCATCGTCGGCCTCAACATGAATGGCTACTTTCACGAGGCC
 TACTTCAAGAAGCACAAAGTTCAACACGGTTCCAGGTGTCCAGCTCAGGAATGTGGACAGTCT
 GAAGAAAGAAGCTTATGAAGTGGAAGTTACAGGGCTGCTCAGTGAGGCTGAGGTTCTGGACC
 ACAGCAAGAACCCTTGTGAAGACTCTTTCCTGCCAGACACAGAGGGCCACACCTACGTGGCC
 TTTATTTCGAATGGAGAAAGATGATGACTTCACCACCTGGACCCAGCTTGCCAAGTGCCCTCA
 TATCTGGGACCTGGATGTGCGTGGCAACCATCGGGGCTGTGGAGATTGTTTCGGAAGAAGA
 ACCACTTCCTGGTGGTGGGGGTCCCGGCTTCCCCCTACTCAGTGAAGAAGCCACCCTCAGTC
 ACCCAATTTTCCTGGAGCCACCCCCAAAGGAGGAGGGAGCCCCAGGAGCCCCAGAACAGAC
ATGAGACCTCCTCCAGGACCCTGCGGGGCTGGGTACTGTGTACCCCCAGGCTGGCTAGCCCT
 TCCCTCCATCCTGTAGGATTTTGTAGATGCTGGTAGGGGCTGGGGCTACCTTGTTTTTAACA
 TGAGACTTAATTACTAACTCCAAGGGGAGGGTTCCCTGCTCCAACACCCCGTTCTGAGTT
 AAAAGTCTATTTATTTACTTCTTGTGGAGAAGGGCAGGAGAGTACCTGGGAATCATTACG
 ATCCCTAGCAGCTCATCCTGCCCTTTGAATACCCTCACTTCCAGGCCTGGCTCAGAATCTA
 ACCTATTTATTGACTGTCTGAGGGCCTTGAAAACAGGCCGAACCTGGAGGGCCTGGATTTT
 TTTTGGGCTGGAATGCTGCCCTGAGGGTGGGGCTGGCTCTTACTCAGGAACTGCTGTGCC
 CAACCCATGGACAGGCCAGCTGGGGCCCACATGCTGACACAGACTCACTCAGAGACCCTTA
 GACTGAGCAGGCTCCTCTCAGCCTTCTCTTTGTCCAGATTTCCAAAGCTGGATAAGTT
 GGTCAATTGATTAAAAAAGGAGAAGCCCTCTGGGAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 54

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA61185

><subunit 1 of 1, 660 aa, 1 stop

><MW: 75220, pI: 6.76, NX(S/T): 0

MDDWKPSPLIKPFGARKKRSWYLTWKYKLTNQRALRRFCQTGAVLFLLVTIVIVNIKLILDTR
RAISEANEDPEPEQDYDEALGRLEPPRRRGSGPRRVLDVEVYSSRSKVYVAVDGTTVLEDEA
REQGRGIHVIVLNQATGHVMAKRVFDTYSPHEDEAMVLFLNMVAPGRVLICTVKDEGSFHLK
DTAKALLRSLGSQAGPALGWRDTWAFVGRKGGPVFGEKHSKSPALSSWGDPVLLKTDVPLSS
AEEAECHWADTELNRRRRRRFC SKVEGYGVCCKDPTPIEFSPDPLPDNKVLNVPVAVIAGN
RPNYLYRMLRSLLSAQGVSPQMITVFIDGYEEPMDEVVALFGLRGIQHTPISIKNARVSQHY
KASLTATFNLFPFAKFAVVLEEDLDIAVDFFSFLSQSIHLLEEDDSLYCISAWNDQGYEHTA
EDPALLYRVETMPGLGWVLRRSLYKEELEPKWPTPEKLWDWDMWMMRMPEQRRGRECIIPDVS
RSYHFGIVGLNMNGYFHEAYFKKKHKFNTVPGVQLRNVD¹SLKKEAYEVEVHRL²LSEAEVLDHS
KNPCEDSFLPDTEGHTYVAFIRMEKDDDFTTWTQLAKCLHIWDL³DVRGNHRGLWRLFRKKNH
FLVVGVPASPYSVKKPPSVTPIFLEPPPKEEGAPGAPEQT

Important features of the protein:

Transmembrane domain:

amino acids 38-55

Homologous region to Mouse GNT1

amino acids 229-660

FIGURE 55

CGGACGCGTGGGCTGCTGGTGGGAAGGCCTAAAGAACTGGAAAGCCCACTCTCTTGGAAACCACCACAC
CTGTTTAAAGAACCTAAGCACCATTAAAGCCACTGGAAATTTGTTGTCTAGTGGTTGTGGGTGAATA
AAGGAGGGCAGAAATGGATGATTTTCATCTCCATTAGCCTGCTGTCTCTGGCTATGTTGGTGGGATGTTA
CGTGGCCCGGAATCATTCCCTTGGCTGTTAATTTCTCAGAGGAACGACTGAAGCTGGTGAAGTGTGTTTGG
GTGCTGGCCTTCTCTGTGGAAGTGTCTGTCAGTCATCGTGCCTGAAGGAGTACATGCCCTTTATGAA
GATATTCTTGAGGGAAAACACCACCAAGCAAGTGAACACATAATGTGATTGCATCAGACAAAGCAGC
AGAAAAATCAGTTGTCCATGAACATGAGCAGCCACGACCACACACAGCTGCATGCCCTATATTGGTG
TTTCCCTCGTTCTGGGCTTCGTTTTTCATGTTGCTGGTGGACCAGATTGGTAACTCCCATGTGCATTCT
ACTGACGATCCAGAAGCAGCAAGGTCTAGCAATTCAAAATCACCACCACGCTGGGTCTGGTTGTCCA
TGCTGCAGCTGATGGTGTGCTTTGGGAGCAGCAGCATCTACTTCACAGACCAGTGTCCAGTTAATTG
TGTTTGTGGCAATCATGCTACATAAGGCACCAGCTGCTTTTGGACTGGTTTCTTCTTGATGCATGCT
GGCTTAGAGCGGAATCGAATCAGAAAGCACTTGTGGTCTTTGCATTGGCAGCACCAGTTATGTCCAT
GGTGACATACTTAGGACTGAGTAAGAGCAGTAAAGAAGCCCTTTCAGAGGTGAACGCCACGGGAGTGG
CCATGCTTTTCTCTGCCGGGACATTTCTTTATGTTGCCACAGTACATGTCCTCCCTGAGGTGGGCGGA
ATAGGGCACAGCCACAAGCCCGATGCCACGGGAGGAGAGGCCTCAGCCGCCTGGAAGTGGCAGCCCT
GGTTCTGGGTTGCCTCATCCCTCTCATCTGTGTCAGTAGGACACCAGCATTAATGTTCAAGGTCCAGC
CTTGGTCCAGGGCCGTTTGGCCATCCAGTGAGAACAGCCGGCACGTGACAGCTACTCACTTCCTCAGTC
TCTTGTCTCACCTTGCGCATCTCTACATGTATTCCTAGAGTCCAGAGGGGAGGTGAGGTTAAACCTG
AGTAATGGAAAAGCTTTTAGAGTAGAAACACATTTACGTTGCAGTTAGCTATAGACATCCCATTTGTGT
TATCTTTTAAAGGCCCTTGACATTTTGCCTTTAATATTTCTCTTAACCCTATTCTCAGGGAAGATG
GAATTTAGTTTAAAGGAAAAGAGGAGAACTTCATACTCACAATGAAATAGTGATTATGAAAATACAGT
GTTCTGTAATTAAGCTATGTCTCTTCTTTCTTAGTTTAGAGGCTCTGCTACTTTATCCATTGATTTT
AACATGGTTCCACCATGTAAGACTGGTGTCTTAGCATCTATGCCACATGCGTTGATGGAAGGTCATA
GCACCCACTCACTTAGATGCTAAAGGTGATTCTAGTTAATCTGGGATTAGGGTCAGGAAAATGATAGC
AAGACACATTGAAAGCTCTCTTTATACTCAAAAGAGATATCCATTGAAAAGGGATGTCTAGAGGGATT
TAAACAGCTCCTTTGGCACGTGCCTCTCTGAATCCAGCCTGCCATTCCATCAAATGGAGCAGGAGAGG
TGGGAGGAGCTTCTAAAGAGGTGACTGGTATTTTGTAGCATTCCTTGTCAAGTTCTCCTTTGCAGAAT
ACCTGTCTCCACATTCCTAGAGAGGAGCCAAGTTCTAGTAGTTTCAGTTCTAGGCTTTCCTTCAAGAA
CAGTCAGATCACAAAGTGTCTTTGGAAATTAAGGGATATTAAATTTAAGTGATTTTTGGATGGTTAT
TGATATCTTTGTAGTAGCTTTTTTTTAAAGACTACCAAATGTATGGTTGTCTTTTTTTTTTTGTTTTT
TTTTTTTTTAAATTATTTCTCTTAGCAGATCAGCAATCCCTCTAGGGACCTAAATACTAGGTCAGCTTT
GGCGACACTGTGTCTTCTCACATAACCACCTGTAGCAAGATGGATCATAAATGAGAAGTGTGTCCTA
TTGATTTAAAGCTTATTGGAATCATGTCTCTTGTCTCTTCGTCTTTTCTTTGCTTTTCTTCTAACTTT
TCCCTCTAGCCTCTCCTCGCCACAATTTGCTGCTTACTGCTGGTGTAAATATTTGTGTGGGATGAATT
CTTATCAGGACAACCACCTTCTCGAACTGTAATAATGAAGATAATAATATCTTTATTCTTTATCCCTT
CAAAGAAATTACCTTTGTGTCAAATGCCGCTTTGTTGAGCCCTTAAATACCACCTCCTCATGTGTAA
ATTGACACAATCACTAATCTGGTAATTTAAACAATTGAGATAGCAAAAGTGTTTAACAGACTAGGATA
ATTTTTTTTTTATATTTGCCAAAATTTTGTAAACCCTGTCTTGTCAAATAAGTGTATAATATTGTAT
TATTAATTTATTTTACTTTCTATACCATTTCAAAACACATTACACTAAGGGGGAACCAAGACTAGTT
TCTTCAGGGCAGTGGACGTAGTAGTTTGTAAAAACGTTTTCTATGACGCATAAGCTAGCATGCCTATG
ATTTATTTTCTTCATGAATTTGTCACTGGATCAGCAGCTGTGGAAATAAAGCTTGTGAGCCCTCTGCT
GGCCACAGTGAGGAAAGTAGCACAAATAGGATACAGTTGTATGTAGTCATTGGCAACAATTGCATACA
ATTTTACTACCAAGAGAAGGTATAGTATGGAAAGTCCAAATGACTTCCTTGATTGGATGTTAACAGCT
GACTGGTGTGAGACTTGAGGTTTCATCTAGTCCTTCAAACTATATGGTTGCCTAGATTCTCTCTGGA
AACTGACTTTGTCAAATAAATAGCAGATTGTAGTGTCAAAAAAA

FIGURE 56

MDDFISISLLSLAMLVGCYVAGIIPLAVNFSEERLKLVTVLGAGLLCGTALAVIVPEGVHAL
YEDILEGKHHQASETHNVIASDKAAEKSVVHEHEHSHDHTQLHAYIGVSLVLGFVFMLLVDQ
IGNSHVHSTDDPEAARSSNSKITTTLGLVVHAAADGVALGAAASTSQTTSVQLIVFVAIMLHK
APAAFGLVSFLMHAGLERNRIRKHLVLFALAAPVMSMVTYLGLSKSSKEALSEVNATGVAML
FSAGTFLYVATVHVLPVGGIGHSHKPDATGGRGLSRLEVAALVLGCLIPLIILSVGHQH

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 37-56, 106-122, 211-230, 240-260, 288-304

FIGURE 57

GCTCGAGGCCGGCGGCGGGAGAGCGACCCGGGCGGCCTCGTAGCGGGGCCCCGGATCCC
CGAGTGGCGGCCGGAGCCTCGAAAAGAGATTCTCAGCGCTGATTTTGAGATGATGGGCTTGG
GAAACGGGCGTCGCAGCATGAAGTCGCCGCCCCCTCGTGCTGGCCGCCCTGGTGGCCTGCATC
ATCGTCTTGGGCTTCAACTACTGGATTGCGAGCTCCCGGAGCGTGGACCTCCAGACACGGAT
CATGGAGCTGGAAGGCAGGGTCCGCAGGGCGGCTGCAGAGAGAGGGCGCCGTGGAGCTGAAGA
AGAACGAGTTCCAGGGAGAGCTGGAGAAGCAGCGGGAGCAGCTTGACAAAATCCAGTCCAGC
CACAACCTTCCAGCTGGAGAGCGTCAACAAGCTGTACCAGGACGAAAAGGCGGTTTTTGGTGAA
TAACATCACACAGGTGAGAGGCTCATCCGAGTGCTGCAAGACCAGTTAAAGACCCTGCAGA
GGAATTACGGCAGGCTGCAGCAGGATGTCTCCAGTTTCAGAAGAACCAGACCAACCTGGAG
AGGAAGTTCTCCTACGACCTGAGCCAGTGCATCAATCAGATGAAGGAGGTGAAGGAACAGTG
TGAGGAGCGAATAGAAGAGGTCACCAAAAAGGGGAATGAAGCTGTAGCTTCCAGAGACCTGA
GTGAAAACAACGACCAGAGACAGCAGCTCCAAGCCCTCAGTGAGCCTCAGCCCAGGCTGCAG
GCAGCAGGCCTGCCACACACAGAGGTGCCACAAGGGAAGGGAACGTGCTTGGTAACAGCAA
GTCCCAGACACCAGCCCCCAGTTCCGAAGTGTTTTGGATTCAAAGAGACAAGTTGAGAAAG
AGGAAACCAATGAGATCCAGGTGGTGAATGAGGAGCCTCAGAGGGACAGGCTGCCGCAGGAG
CCAGGCCGGGAGCAGGTGGTGGGAAGACAGACCTGTAGGTGGAAGAGGCTTCGGGGGAGCCGG
AGAACTGGGCCAGACCCACAGGTGCAGGCTGCCCTGTGAGTGAGCCAGGAAAATCCAGAGA
TGGAGGGCCCTGAGCGAGACCAGCTTGTATCCCCGACGGACAGGAGGAGGAGCAGGAAGCT
GCCGGGGAAGGGAGAAACCAGCAGAACTGAGAGGAGAAGATGACTACAACATGGATGAAAA
TGAAGCAGAACTCTGAGACAGACAAGCAAGCAGCCCTGGCAGGGAATGACAGAAACATAGATG
TTTTTAATGTTGAAGATCAGAAAAGAGACACCATAAAATTTACTTGATCAGCGTGAAAAGCGG
AATCATACACTCTGAATTGAACTGGAATCACATATTTTCAACAGGGCCGAAGAGATGACTA
TAAAATGTTTCATGAGGGACTGAATACTGAAAACGTGAAATGTACTAAATAAAATGTACATCTGA

FIGURE 58

MMGLGNRRSMKSPPLVLAALVACIIIVLGFNYWIIASSRSVDLQTRIMELEGRVRRAAAERGA
VELKKNEFQGELEKQREQLDKIQSSHNFQLESVNKLYQDEKAVLVNNITTGERLIRVLQDQL
KTLQARNYGRLQQDVLQFQKNQTNLERKFSYDLSQCINQMKEVKEQCEERIEEVTKKGNEAVA
SRDLSENNDQRQQQLQALSEPQPRQLQAAGLPHTTEVPQGKGNVLGNSKSQTPAPSSEVVLDSKR
QVEKEETNEIQVVNEEPQRDRLPQEPGREQVVEDRPVGGRGFGGAGELGQTPQVQAALSVSQ
ENPEMEGPERDQLVIPDGQEEEQEAAGEGRNQQLRGEDDYNMDENEAESETDKQAALAGND
RNIDVFNVEDQKRDTINLLDQREKRNHTL

Signal peptide:

amino acids 1-29

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 59

GGATGCAGAAAGCCTCAGTGTGTGCTCTTCTGGCCTGGGTCTGCTTCTCTTCTACGCTGGCATTGCCCTCTTCA
 CCAGTGGCTTCTGCTCACCCTTTGGAGCTCACCACCATAGCAGCTGCCAAGAGCCCCAGGCCCTGGGTCCC
 TGCCATGGGGGAGCCAAGGGAAACCTGGGGCCTGCTGGATGGCTTCCCGATTTTCGCGGGTTGTGTTGGTGCTGA
 TAGATGCTCTGCGATTTGACTTCGCCCAGCCCCAGCATTACACGTGCCTAGAGAGCCTCCTGTCTCCCTACCCT
 TCCTGGGCAAACCTAAGCTCCTTGAGAGGATCCTGGAGATTAGCCCCACCATGCCCGCTCTACCGATCTCAGG
 TTGACCCTCCTACCACCACCATGCAGCGCTCAAGGCCCTCACCCTGGCTCACTGCCTACCTTTATTGATGCTG
 GTAGTAACCTCGCCAGCCACGCCATAGTGGAAAGACAATCTCATTAAAGCAGCTCACCAGTGCAGGAAGGCGTGTAG
 TCTTCATGGGAGATGATACCTGGAAAAGACCTTTTCCCTGGTGCTTTCTCCAAAGCTTTCTTCTTCCCATCCTTCA
 ATGTGAGAGACCTAGACACAGTGGACAATGGCATCCTGGAAACACCTCTACCCACCATGGACAGTGGTGAATGGG
 ACGTGCTGATTGCTCACTTCTGGGTGTGGACCACTGTGGCCACAAGCATGGCCCTCACCACCCTGAAATGGCCA
 AGAAACTTAGCCAGATGGACCAGGTGATCCAGGGACTTGTGGAGCGTCTGGAGAATGACACACTGCTGGTAGTGG
 CTGGGGACCATGGGATGACCACAAATGGAGACCATGGAGGGGACAGTGAGCTGGAGGTCTCAGCTGCTCTCTTTC
 TGTATAGCCCCACAGCAGTCTTCCCCAGCACCCACAGAGGAGCCAGAGGTGATTCTCAAGTTAGCCTTGTGC
 CCACGCTGGCCCTGCTGCTGGGCCCTGCCATCCCATTTGGGAATATCGGGGAAGTGATGGCTGAGCTATTCTCAG
 GGGGTGAGGACTCCCAGCCCCACTCCTCTGCTTTAGCCCCAAGCCTCAGCTCTCCATCTCAATGCTCAGCAGGTGT
 CCCGATTTCTTCATACCTACTCAGCTGCTACTCAGGACCTTCAAGCTAAGGAGCTTCATCAGCTGCAGAACCTCT
 TCTCCAAGGCCCTCTGCTGACTACCAGTGGCTTCTCCAGAGCCCCAAGGGGGCTGAGGCGACACTGCCGACTGTGA
 TTGCTGAGCTGCAGCAGTTCCTGCGGGGAGCTCGGGCCATGTGCATCGAGTCTTGGGCTCGTTTCTCTCTGGTCC
 GCATGGCGGGGGGTACTGCTCTCTTGGCTGCTTCTGCTTTATCTGCCTGCTGGCATCTCAGTGGGCAATATCCC
 CAGGCTTTCCATTCTGCCCTCTACTCCTGACACCTGTGGCCTGGGGCCTGGTTGGGGCCATAGCGTATGCTGGAC
 TCCTGGGAACCTATTGAGCTGAAGCTAGATCTAGTGCTTCTAGGGGCTGTGGCTGCAGTGAGCTCATTCTCCCTT
 TTCTGTGGAAGCCTGGGCTGGCTGGGGGTCCAAGAGGGCCCTGGCAACCCTGTTTCCCATCCCTGGGCCCCGTCC
 TGTACTCCTGCTGTTTCGCTTGGCTGTGTTCTTCTCTGATAGTTTGTGTAGCTGAGGCCAGGGCCACCCCT
 TCCTTTTGGGCTCATTATCCTGCTCCTGGTGTCCAGCTTCACTGGGAGGGCCAGCTGCTTCCACCTAAGCTAC
 TCACAATGCCCCGCCTTGGCACTTCAGCCACAACAAACCCCCACGGCACAATGGTGCATATGCCCTGAGGCTTG
 GAATTGGGTGCTTTTATGTACAAGGCTAGCTGGGCTTTTTCATCGTTGCCCTGAAGAGACACCTGTTTGGCACT
 CCTCTCCCTGGCTGAGTCCCTCTGGCATCCATGGTGGGTGGTTCGAGCCAGAATTTATGGTATGGAGCTTGTGTGG
 CGGCGCTGGTGGCCCTGTAGCTGCCGTGCGCTTGTGGCTTCGCGCTATGGTAATCTCAAGAGCCCCGAGCCAC
 CCATGCTCTTTGTGCGCTGGGACTGCCCTAATGGCATTGGGTACTGCTGCCTACTGGGCATTGGCGTCGGGG
 CAGATGAGGCTCCCCCGCTCTCCGGTCTGGTCTTGGGGCATCCATGGTGTGCTCGGCTGTAGCAGGGC
 TGGCTGCTTCAGGGCTCGCGCTGCTGCTCTGGAAGCCTGTGACAGTGTGGTGAAGGCTGGGGCAGGCGCTCCAA
 GGACCAGGACTGTCTCACTCCCTTCTCAGGCCCCCCACTTCTCAAGCTGACTTGGATTATGTGGTCCCTCAAA
 TCTACCGACACATGCAGGAGGAGTTCGGGGGCGGTTAGAGAGGACCAAATCTCAGGGTCCCCTGACTGTGGCTG
 CTTATCAGTTGGGGAGTGTCTACTCAGCTGCTATGGTCACAGCCCTCACCCTGTTGGCCTTCCCCTTCTGCTGT
 TGCATGCGGAGCGCATCAGCCTTGTGTTCTGCTTCTGTTTCTGCAGAGCTTCTTCTCTACATCTGCTTGGCTG
 CTGGGATACCCGTACCACCCCTGGTCTTTTACTGTGCCATGGCAGGCAGTCTCGGCTTGGGCCCTCATGGCCA
 CACAGACCTTCTACTCCACAGGCCACCAGCCTGTCTTCCAGCCATCCATTGGCATGCAGCCTTCGTGGGATTCC
 CAGAGGGTCATGGCTCCTGTACTTGGCTGCCTGCTTTGCTAGTGGGAGCCAACACCTTTGCCTCCCACCTCCTCT
 TTGCAGTAGGTTGCCCACTGCTCCTGCTCTGGCCTTTCTGTGTGAGAGTCAAGGGCTGCGGAAGAGACAGCAGC
 CCCCAGGGAATGAAGCTGATGCCAGAGTCAGACCCGAGGAGGAAGAGGAGCCACTGATGGAGATGCGGCTCCGGG
 ATGCGCCTCAGCACTTCTATGCAGCACTGCTGCAGCTGGGCCTCAAGTACCTCTTTATCCTTGGTATTTCAGATTCT
 TGGCCTGTGCCTTGGCAGCCTCCATCCTTCGAGGCATCTCATGGTCTGGAAAGTGTGTTGCCCTAAGTTTCATAT
 TTGAGGCTGTGGGCTTCATTGTGAGCAGCGTGGGACTTCTCCTGGGCATAGCTTTGGTGATGAGAGTGGATGGTG
 CTGTGAGCTCCTGGTTAGGCAGCTATTTCTGGCCCAGCAGAGG**TAG**CCTAGTCTGTGATTACTGGCACTTGGCT
 ACAGAGAGTGTGGAGAACAGTGTAGCCTGGCCTGTACAGGTACTGGATGATCTGCAAGACAGGCTCAGCCATAC
 TCTTACTATCATGCAGCCAGGGGCCGCTGACATCTAGGACTTCATTATTCTATAATTAGGACCACAGTGGAGTA
 TGATCCCTAACTCCTGATTTGGATGCATCTGAGGGACAAGGGGGGCGGTCTCCGAAGTGGAATAAAATAGGCCGG
 GCGTGGTGACTTGCACCTATAATCCCAGCACTTTGGGAGGCAGAGGTGGGAGGATTGCTTGGTCCCAGGAGTTCA
 AGACCAGCCTGTGGAACATAACAAGACCCCGTCTCTACTATTTAAAAAAAGTGAATAAAATGATAATAT

FIGURE 60

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62809

<subunit 1 of 1, 1089 aa, 1 stop

<MW: 118699, pI: 8.49, NX(S/T): 2

MQKASVLLFLAWVCFLFYAGIALFTSGFLLTRLELTNHSSCQEPPGPGSLPWGSQGKPGACW
 MASRFSRVVLVLIDALRFDFAPQHQSHVPREPPVSLPFLGKLSSLQRILEIQPHHARLYRSQ
 VDPPTTTMQRLKALTGSLPTFIDAGSNFASHAIVEDNLIKQLTSAGRRVVFMGDDTWKDLF
 PGAFSKAFFFPFNFVRDLDTVDNGILEHLYPTMDSGEWDVLIHFLGVDHCGHKHGPHPHP
 AKKLSQMDQVIQGLVERLENDTLLVVAGDHGMTTNGDHGGDSELEVSAALFLYSPTAVFPST
 PPEEPEVIPQVSLVPTLALLLGLPIPFNGIENVMAELFSGGEDSQPHSSALAQAASALHLNAQ
 QVSRFLHTYSAATQDLQAKELHQLQNLFSKASADYQWLLQSPKGAEATLPTVIAELQQFLRG
 ARAMCIESWARFSLVRMAGGTALLAASCFCILLASQWAIISPGFPFCPLLLTPVAWGLVGAIA
 YAGLLGTIELKLDLVLLGAVAAVSSFLPFLWKAWAGWGSKRPLATLFPIPGPVLLLLLFRLA
 VFFSDSFVVAEARATPFLGSGFILLVVLHWEQQLLPKLLTMPRLGTSATTNPPRHNGAY
 ALRLGIGLLLCTRLAGLFHRCPEETPVCHSSPWLSPLASMVGGRAKNLWYGACVAALVALLA
 AVRLWLRRYGNLKSPEPPMLFVRWGLPLMALGTAAYWALASGADEAPPRLRVLVSGASMLP
 RAVAGLAASGLALLLWKPVTVLVKAGAGAPRTRTVLTPFSGPPTSQADLDYVVPQIYRHMQE
 EFRGRLERTKSQGPLTVAAYQLGVSYSAMVTALTLLAFPLLLLHAERISLVFLLLFLQSFL
 LLHLLAAGIPVTTGPFPTVPWQAVSAWALMATQTFYSTGHQPVFPAIHWHAFAVGFPEGHGS
 CTWLPALLVGANTFASHLLFAVGCPLLLLPFLCESQGLRKRQQPPGNEADARVRPEEEEEEP
 LMEMRLRDAPQHFYAALLQLGLKYLFIILGIQILACALAASILRRHLMVWKVFAPKFIFEAVG
 FIVSSVGLLLGIALVMRVDGAVSSWFRQLFLAQQR

Important features:**Signal peptide:**

amino acids 1-16

Transmembrane domains:

amino acids 317-341, 451-470, 481-500, 510-527, 538-555, 831-850,
 1016-1034, 1052-1070

Leucine zipper pattern.

amino acids 843-864

N-glycosylation sites.

amino acids 37-40, 268-271

FIGURE 61

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGT
 GTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA**ATGT**
 CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATT
 TCATCACAAATTGGCCCAACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGT
 CCTGACAGCTCCAGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAA
 TATACTCCAATCTGAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCC
 CAGTGTGTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGT
 ACACGTGGAGTCCTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTG
 CCAGGACTTTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTG
 CCCATATCTATTACCGTGTTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCA
 CGTTGGCAAAGAGAAACACCCAGCAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAA
 GATTCTTTGTGCTGCTGAAAAAATCGTGATTAACCTTTATCACCTCAATATCTCGGATGAT
 TCTAAAATTTCTCATCAGGATATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAA
 TGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAG
 GGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCT
 CTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAAACAGTCATTGAATATGA
 ATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGG
 AGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCG
 CAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCTGGCGCAGGAGCA
 CACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGATCCCC
 AACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAG
 CCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGC
 TCCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGT
 TATATGTGCAGATGGAAAAC**TGA**TGCCAACACTTCCTTTTGCCTTTTGTTTCCTGTGCAAC
 AAGTGAGTCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGT
 TTGTCAGTGTCTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTG
 GTTCATGCATGTAGGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGT
 TGTTCTATGCAGAGAAAGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGG
 TGGGTGT

FIGURE 62

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62815

<subunit 1 of 1, 442 aa, 1 stop

<MW: 49932, pI: 4.55, NX(S/T): 5

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQ
QIYSNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQ
CARTLKDQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDF
KRFFVPAEKIVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKH
LGYASHLMEIFCDSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDCAGPEEQELSL
QEEVSTQGTLLESQAALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEEPSTTLVDWD
PQTGRLCIPSLSSFDQDSEGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEW
GLYVQMEN

Important features:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 63

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTCTCTGCGGGGAGACGCCAGCCTGCG
TCTGCC**ATG**GGGCTCGGGTTGAGGGGCTGGGGACGTCTCTGCTGACTGTGGCCACCGCCCT
GATGCTGCCCCTGAAGCCCCCGCAGGCTCCTGGGGGGGCCAGATCATCGGGGGCCACGAGG
TGACCCCCCACTCCAGGCCCTACATGGCATCCGTGCGCTTCGGGGGCCAACATCACTGCGGA
GGCTTCCTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCCACAGAGACCT
CCGCACTGGCCTGGTGGTGGTGGGCGCCACGTCTGAGTACTGCGGAGCCCAACCCAGCAGG
TGTTTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCCAACGAC
ATCTGCCTGCTGCGGCTGAACGGCTCTGCTGTCCTGGGCCCTGCAGTGGGGCTGCTGAGGCT
GCCAGGGAGAAGGGCCAGGCCCCCACAGCGGGGACACGGTGCCGGGTGGCTGGCTGGGGCT
TCGTGTCTGACTTTGAGGAGCTGCCGCTGGACTGATGGAGGCCAAGGTCCGAGTGCTGGAC
CCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATGCTCTGCACCCGCAG
TGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGGGCCCTGGTGTGCAGGA
ACCGGGCTCACGGCCTCGTTTCCTTCTCGGGCCTCTGGTGCGGCGACCCCAAGACCCCCGAC
GTGTACACGCAGGTGTCCGCCTTTGTGGCCTGGATCTGGGACGTGGTTCGGCGGAGCAGTCC
CCAGCCCGGCCCCCTGCCTGGGACCACCAGGCCCCCAGGAGAAGCCGCC**TGA**GCCACAACCT
TGCGGCATGCAAATGAGATGGCCGCTCCAGGCCTGGAATGTTCCGTGGCTGGGCCCCACGGG
AAGCCTGATGTTCAAGGTTGGGGTGGGACGGGCAGCGGTGGGGCACACCCATTCCACATGCA
AAGGGCAGAAGCAAACCCAGTAAAATGTAACTGACAAAAAAAAAAAAAAAAAAAAAGAAA

FIGURE 64

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62845

><subunit 1 of 1, 283 aa, 1 stop

><MW: 30350, pI: 9.66, NX(S/T): 2

MGLGLRGWGRPLLTVATALMLPVKPPAGSWGAIIGGHEVTPHSRPYMASVRFGGQHHCGGF
LLRARWVVSAAHCFSHRDLRTGLVVLGAHVLSTAEPTQQVFGIDALTTHPDYHPMTHANDIC
LLRLNGSAVLGPAVGLLRLPGRRARPPTAGTRCRVAGWGFVSDFEELPPGLMEAKVRVLDPD
VCNSSWKGHLLTLMCTRSGDSHRRGFCSADSGGPLVCRNRAHGLVSEFSGLWCGDPKTPDVY
TQVSAFVAWIWDVVRSSPQPGPLPGTTRPPGEAA

Signal peptide:

amino acids 1-30

60
50
40
30
20
10
0

FIGURE 65

GAGCTACCCAGGCGGCTGGTGTGCAGCAAGCTCCGCGCCGACTCCGGACGCCTGACGCCTGA
 CGCCTGTCCCCGGCCCGGC**ATG**AGCCGCTACCTGCTGCCGCTGTCGGCGCTGGGCACGGTAG
 CAGGCGCCGCGTGTCTCAAGGACTATGTACCCGGTGGGGCTTGCCCCAGCAAGGCCACC
 ATCCCTGGGAAGACGGTCATCGTGACGGGCGCCAACACAGGCATCGGGAAGCAGACCGCCTT
 GGAAGTGGCCAGGAGAGGAGGCAACATCATCCTGGCCTGCCGAGACATGGAGAAGTGTGAGG
 CGGCAGCAAAGGACATCCGCGGGGAGACCCTCAATCACCATGTCAACGCCCCGGCACCTGGAC
 TTGGCTTCCCTCAAGTCTATCCGAGAGTTTGCAGCAAAGATCATTGAAGAGGAGGAGCGAGT
 GGACATTCTAATCAACAACGCGGGTGTGATGCGGTGCCCCCACTGGACCACCGAGGACGGCT
 TCGAGATGCAGTTTGGCGTTAACCACCTGGGTCACTTTCTCTTGACAAACTTGCTGCTGGAC
 AAGCTGAAAGCCTCAGCCCCCTTCGCGGATCATCAACCTCTCGTCCCTGGCCCATGTTGCTGG
 GCACATAGACTTTGACGACTTGAAGTGGCAGACGAGGAAGTATAACACCAAAGCCGCCTACT
 GCCAGAGCAAGCTCGCCATCGTCCTCTTACCAAGGAGCTGAGCCGGCGGCTGCAAGGCTCT
 GGTGTGACTGTCAACGCCCTGCACCCCGGCGTGGCCAGGACAGAGCTGGGCAGACACACGGG
 CATCCATGGCTCCACCTTCTCCAGCACCACTCGGGCCCCTCTTCTGGCTGCTGGTCAAGA
 GCCCCGAGCTGGCCGCCAGCCAGCACATACTGGCCGTGGCGGAGGAACTGGCGGATGTT
 TCCGGAAAGTACTTCGATGGACTCAAACAGAAGGCCCCCGCCCCGAGGCTGAGGATGAGGA
 GGTGGCCCGGAGGCTTTGGGCTGAAAGTGGCCGCTGGTGGGCTTAGAGGCTCCCTCTGTGA
 GGGAGCAGCCCCCTCCCCAG**TAA**CCCTCTGGAGCAGATTTGAAAGCCAGGATGGCGCCTCCAG
 ACCGAGGACAGCTGTCCGCCATGCCCGCAGCTTCTGGCACTACCTGAGCCGGGAGACCCAG
 GACTGGCGGCCGCCATGCCCGCAGTAGGTTCTAGGGGGCGGTGCTGGCCGCAGTGGACTGGC
 CTGCAGGTGAGCACTGCCCCGGGCTCTGGCTGGTTCCGTCTGCTCTGCTGCCAGCAGGGGAG
 AGGGGCCATCTGATGCTTCCCCTGGGAATCTAAACTGGGAATGGCCGAGGAGGAAGGGGCTC
 TGTGCACTTGCAGGCCACGTCAGGAGAGCCAGCGGTGCCTGTCGGGGAGGGTTCCAAGGTGC
 TCCGTGAAGAGCATGGGCAAGTTGTCTGACACTTGGTGGATTCTTGGGTCCCTGTGGGACCT
 TGTGCATGCATGGTCCTCTCTGAGCCTTGTTTTCTTCAGCAGTGAGATGCTCAGAATAACTG
 CTGTCTCCCATGATGGTGTGGTACAGCGAGCTGTTGTCTGGCTATGGCATGGCTGTGCCGGG
 GGTGTTTGCTGAGGGCTTCCGTGTCCAGAGCCAGCCAGAGAGCAGGTGCAGGTGTCATCCC
 GAGTTCAGGCTCTGCACGGCATGGAGTGGGAACCCACAGCTGCTGCTACAGGACCTGGGA
 TTGCTGGGACTCCCACCTTCCTATCAATTCTCATGGTAGTCCAAACTGCAGACTCTCAAAC
 TTGCTCATTT

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64842

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35932, pI: 8.45, NX(S/T): 1

MSRYLLPLSALGTVAGAAVLLKDYVTGGACPSKATIPGKTVIVTGANTGIGKQTALELARRG
GNIILACRDMEKCEAAAKDIRGETLNHHVNARHLDLASLKSIREFAAKIIEEEERVLDILINN
AGVMRCPHWTTEDGFEMQFGVNHLLGHFLLTNLLLDKLGKASAPSRIINLSSLAHVAGHIDFDD
LWQTRKYNTKAAAYCQSKLAIVLFTKELSRRLQGSGVTVNALHPGVARTELGRHTGIHGSTF
SSTTLGPFIWLLVKSPELAAQPSTYLAVAEELADVSGKYFDGLKQKAPAPEAEDEEVARRLW
AESARLVGLEAPSVREQPLPR

Signal peptide:

amino acids 1-17

FIGURE 67

GAAGTTCGCGAGCGCTGGC**ATG**TGGTCCTGGGGCGCGGCTGGCGGCGCTGCTGGCGGTGCTG
 GCGCTCGGGACAGGAGACCCAGAAAGGGCTGCGGCTCGGGGCGACACGTTCTCGGCGCTGAC
 CAGCGTGGCGCGCGCCCTGGCGCCCGAGCGCCGGCTGCTGGGGCTGCTGAGGCGGTACCTGC
 GCGGGGAGGAGGCGCGGCTGCGGGACCTGACTAGATTCTACGACAAGGTACTTTCTTTGCA
 GAGGATTCAACAACCCCTGTGGCTAACCCCTCTGCTTGCATTTACTCTCATCAAACGCCTGCA
 GTCTGACTGGAGGAATGTGGTACATAGTCTGGAGGCCAGTGAGAACATCCGAGCTCTGAAGG
 ATGGCTATGAGAAGGTGGAGCAAGACCTTCCAGCCTTTGAGGACCTTGAGGGAGCAGCAAGG
 GCCCTGATGCGGCTGCAGGACGTGTACATGCTCAATGTGAAAGGCCTGGCCCCGAGGTGTCTT
 TCAGAGAGTCACTGGCTCTGCCATCACTGACCTGTACAGCCCCAAACGGCTCTTTTCTCTCA
 CAGGGGATGACTGCTTCCAAGTTGGCAAGGTGGCCTATGACATGGGGGATTATTACCATGCC
 ATTCCATGGCTGGAGGAGGCTGTCACTCTTCCGAGGATCTTACGGAGAGTGGAAGACAGA
 GGATGAGGCAAGTCTAGAAGATGCCTTGGATCACTTGGCCTTTGCTTATTTCCGGGCAGGAA
 ATGTTTCGTGTGCCCTCAGCCTCTCTCGGGAGTTTCTTCTCTACAGCCCAGATAATAAGAGG
 ATGGCCAGGAATGTCTTGAAATATGAAAGGCTCTTGGCAGAGAGCCCCAACCCACGTGGTAGC
 TGAGGCTGTCATCCAGAGGCCCAATATACCCACCTGCAGACCAGAGACACCTACGAGGGGC
 TATGTCAGACCCTGGGTTCCAGCCCACTCTCTACCAGATCCCTAGCCTCTACTGTTCTCTAT
 GAGACCAATTCCAACGCCTACCTGCTGCTCCAGCCCATCCGGAAGGAGGTCATCCACCTGGA
 GCCCTACATTGCTCTCTACCATGACTTCGTCACTGACTCAGAGGCTCAGAAAATTAGAGAAC
 TTGCAGAACCATTGGCTACAGAGGTCACTGGTGGCATCAGGGGAGAAGCAGTTACAAGTGGAG
 TACCGCATCAGCAAAAGTGCCTGGCTGAAGGACACTGTTGACCCAAAACCTGGTGACCCTCAA
 CCACCGCATTGCTGCCCTCACAGGCCTTGATGTCCGGCCTCCCTATGCAGAGTATCTGCAGG
 TGGTGAACATATGGCATCGGAGGACACTATGAGCCTCACTTTGACCATGCTACGTCACCAAGC
 AGCCCCCTCTACAGAATGAAGTCAGGAAACCGAGTTGCAACATTTATGATCTATCTGAGCTC
 GGTGGAAGCTGGAGGAGCCACAGCCTTCATCTATGCCAACCTCAGCGTGCCTGTGGTTAGGA
 ATGCAGCACTGTTTTGGTGGAACTGCACAGGAGTGGTGAAGGGGACAGTGACACACTTCAT
 GCTGGCTGTCCTGTCCTGGTGGGAGATAAGTGGGTGGCCAACAAGTGGATACATGAGTATGG
 ACAGGAATTCCGCAGACCCTGCAGCTCCAGCCCTGAAGACT**TGA**ACTGTTGGCAGAGAGAAGC
 TGGTGGAGTCCCTGTGGCTTTCCAGAGAAGCCAGGAGCCAAAAGCTGGGGTAGGAGAGGAGAA
 AGCAGAGCAGCCTCCTGGAAGAAGGCCTTGTCAGCTTTGTCTGTGCCTCGCAAATCAGAGGC
 AAGGGAGAGGTTGTTACCAGGGGACACTGAGAATGTACATTTGATCTGCCCCAGCCACGGAA
 GTCAGAGTAGGATGCACAGTACAAAGGAGGGGGGAGTGGAGGCCTGAGAGGGAAGTTTCTGG
 AGTTCAGATACTCTCTGTTGGGAACAGGACATCTCAACAGTCTCAGGTTTCGATCAGTGGGTC
 TTTTGGCACTTTGAACCTTGACCACAGGGACCAAGAAGTGGCAATGAGGACACCTGCAGGAG
 GGGCTAGCCTGACTCCCAGAACTTTAAGACTTTCTCCCCACTGCCTTCTGCTGCAGCCCAAG
 CAGGGAGTGTCCCCCTCCCAGAAAGCATATCCCAGATGAGTGGTACATTATATAAGGATTTTT
 TTTAAGTTGAAAACAACCTTTCTTTTCTTTTGTATGATGGTTTTTTAACACAGTCATTAAAA
 ATGTTTATAAATCAAAA

FIGURE 68

MGPGARLAALLAVLALGTGDPERAAARGDTFSALTSVARALAPERRLGLLRRYLRGEEARL
RDLTRFYDKVLSLHEDSTTPVANPLLAFTLIKRLQSDWRNVVHSLEASENIRALKDGYEKVE
QDLPAFEDLEGAARALMRLQDVYMLNVKGLARGVFQRVTGSAITDLYSPKRLFSLTGDDCFQ
VGKVAYDMGDYYHAIPWLEEAVSLFRGSYGEWKTEDEASLEDALDHLAFAYFRAGNVSCALS
LSREFLLYSPDNKRMARNVLKYERLLAESPNHVVAEAVIQRPNIHPLQTRDTYEGLCQTLGS
QPTLYQIPSLYCSYETNSNAYLLLQPIRKEVIHLEPYIALYHDFVSDSEAQKIRELAEPWLQ
RSVVASGEKQLQVEYRISKSAWLKDTVDPKLVTLNHRIAALTGLDVRPPYAEYLQVVNYGIG
GHYEPHFDHATSPSSPLYRMKSGNRVATFMIYLSSVEAGGATAFIYANLSVPVVRNAALFWW
NLHRSGEGDSDTLHAGCPVLVGDKWVANKWIHEYGQEFRRPCSSSPED

Signal peptide:

amino acids 1-19

FIGURE 69

GAGATAGGGAGTCTGGGTTTAAGTTCCTGCTCCATCTCAGGAGCCCCTGCTCCCACCCCTAG
 GAAGCCACCAGACTCCACGGTGTGGGGCCAATCAGGTGGAATCGGCCCTGGCAGGTGGGGCC
 ACGAGCGCTGGCTGAGGGACCGAGCCGGAGAGCCCCGGAGCCCCCGTAACCCGCGCGGGGAG
 CGCCAGG**ATG**CCGCGCGGGGACTCGGAGCAGGTGCGCTACTGCGCGCGCTTCTCCTACCTC
 TGGCTCAAGTTTTCACTTATCATCTATTCCACCGTGTTCTGGCTGATTGGGGCCCTGGTCCT
 GTCTGTGGGCATCTATGCAGAGGTTGAGCGGCAGAAATATAAAACCCTTGAAAGTGCCTTCC
 TGGCTCCAGCCATCATCCTCATCCTCCTGGGCGTCGTCATGTTTCATGGTCTCCTTCATTGGT
 GTGCTGGCGTCCCTCCGTGACAACCTGTACCTTCTCCAAGCATTTCATGTACATCCTTGGGAT
 CTGCCTCATCATGGAGCTCATTGGTGGCGTGGTGGCCTTGACCTTCCGGAACCAGACCATTG
 ACTTCCTGAACGACAACATTTCGAAGAGGAATTGAGAACTACTATGATGATCTGGACTTCAAA
 AACATCATGGACTTTGTTTCAGAAAAAGTTCAAGTGCTGTGGCGGGGAGGACTACCGAGATTG
 GAGCAAGAATCAGTACCACGACTGCAGTGCCCCCTGGACCCCTGGCCTGTGGGGTGCCCTACA
 CCTGCTGCATCAGGAACACGACAGAAGTTGTCAACACCATGTGTGGCTACAAAACATCGAC
 AAGGAGCGTTTTAGTGTGCAGGATGTATCTACGTGCGGGGCTGCACCAACGCCGTGATCAT
 CTGGTTCATGGACAACTACACCATCATGGCGTGCATCCTCCTGGGCATCCTGCTTCCCCAGT
 TCCTGGGGGTGCTGCTGACGCTGCTGTACATCACCCGGGTGGAGGACATCATCATGGAGCAC
 TCTGTCACTGATGGGCTCCTGGGGCCCCGGTGCCAAGCCCAGCGTGGAGGCGGCAGGCACGGG
 ATGCTGCTTGTGCTACCCCAAT**TAG**GGGCCCAGCCTGCCATGGCAGCTCCAACAAGGACCGTC
 TGGGATAGCACCTCTCAGTCAACATCGTGGGGCTGGACAGGGCTGCGGCCCCCTCTGCCCCA
 CTCAGTACTGACCAAAGCCAGGGCTGTGTGTGCCTGTGTGTAGGTCCCACGGCCTCTGCCTC
 CCCAGGGAGCAGAGCCTGGGCCTCCCCTAAGAGGCTTTCCCCGAGGCAGCTCTGGAATCTGT
 GCCACCTGGGGCCTGGGGAACAAGGCCCTCCTTTCTCCAGGCCTGGGCTACAGGGGAGGGA
 GAGCCTGAGGCTCTGCTCAGGGCCCATTTCATCTCTGGCAGTGCCTTGGCGGTGGTATTCAA
 GGCAGTTTTGTAGCACCTGTAATTGGGGAGAGGGAGTGTGCCCTCGGGGCAGGAGGGAAGG
 GCATCTGGGGAAGGGCAGGAGGGAAGAGCTGTCCATGCAGCCACGCCCATGGCCAGGTTGGC
 CTCTTCTCAGCCTCCCAGGTGCCTTGAGCCCTCTTGCAAGGGCGGCTGCTTCCTTGAGCCTA
 GTTTTTTTTTTACGTGATTTTTGTAAACATTCATTTTTTTGTACAGATAACAGGAGTTTCTGAC
 TAATCAAAGCTGGTATTTCCCCGCATGTCTTATTCTTGCCCTTCCCCCAACCAGTTTGTAA
 TCAAACAATAAAAAACATGTTTTGTTTTGTTTTTAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64863

><subunit 1 of 1, 294 aa, 1 stop

><MW: 33211, pI: 5.35, NX(S/T): 3

MPRGDSEQVRYCARFSYLWLKFSLLIYSTVFWLIGALVLSVGIYAEVERQKYKTLES AFLAP
AIILILLGVVMFMVSFIGVLASLRDNL YLLQAFMYILGICLIMELIGGVVALTFRNQ TIDFL
NDNIRRG IENYYDDLDFKNIMDFVQKKFKCCGGEDYRDWSKNQYHDCSAPGPLACGVPYTCC
IRNTTEVVNTMCGYKTIDKERFSVQDVIIYVRGCTNAV IIFMDNYTIMACILLGILLPQFLG
VLLTLLYITRVEDIIMEHSVTDG LLGPGAKPSVEAAGTGCCLCYPN

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 22-42, 57-85, 93-116, 230-257

FIGURE 71

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGA
CACCTGGGAAG**ATG**GCCGGCCCGTGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACC
TTGATCCAAGCCACCCTCAGTCCCACTGCAGTTCTCATCCTCGGCCCAAAGTCATCAAAGA
AAAGCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGC
TCAGTGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACC
GTCCTGAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAA
GCCCTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCA
ACACGCCCCCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATC
CGCATGGACACCAGTGCAAGTGGCCCCACCCGCCTGGTCCTCAGTGACTGTGCCACCAGCCA
TGGGAGCCTGCGCATCCAAGTGTGTATAAGCTCTCCTTCCTGGTGAACGCCTTAGCTAAGC
AGGTCATGAACCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTGTCCCGTG
ATCGAGGCTTCCTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCCATTTTC
CCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATTTC
AGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCAATAAC
TCTGCAGCTTCCCTGACAATGCCCACCCTGGACAACATCCCGTTTCAGCCTCATCGTGAGTCA
GGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGAATTTCATGGTCCTGTTGG
ACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAG
GCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTT
TTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTTTCCCTCCA
GTGAAGCCCTCCGCCCTTTGTTTACCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTAC
ACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGAT
GAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCC
ACTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTG
AAGGCCTTGGGATTCGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCC
AGCCTCCTTGTGGAAACCCAGCTCTCCTGTCTCCCAG**TGA**AGACTTGGATGGCAGCCATCAG
GGAAGGCTGGGTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAAT
CAATAAACACTTGCCTGTGAAAAA

FIGURE 72

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64881

><subunit 1 of 1, 484 aa, 1 stop

><MW: 52468, pI: 7.14, NX(S/T): 3

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQPLLSAM
REKPAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPL
VKTIVEFHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMN
LLVPSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYL
GAKLLDSQGKVTWKFNNSAASLTMPITLDNIPFSLIVSQDVVKAABAVALSPEEFMVLLDSVL
PESAHRLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEAL
RPLFTLGIEASSEAQFYTKGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSIL
LPNQNGKLRSGVPVSLVKALGFEEAESSLTKDVLTPASLWKPPSPVSQ

Important features of the protein:**Signal peptide:**

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 73

GAGCGAAC**ATG**GACGCGGTTGGCGGTTTTGGTGTGTCTCTGTGACCATGGTGGTGGCGCTG
 CTCATCGTTTTGCGACGTTCCCTCAGCCTCTGCCCAAAGAAAGAAGGAGATGGTGTATCTGA
 AAAGGTTAGTCAGCTGATGGAATGGACTAACAAAAGACCTGTAATAAGAATGAATGGAGACA
 AGTTCCGTCGCCTTGTGAAAGCCCCACCGAGAAATTACTCCGTTATCGTCATGTTCACTGCT
 CTCCAAGTGCATAGACAGTGTGTCTTTGCAAGCAAGCTGATGAAGAATTCCAGATCCTGGC
 AAACCTCCTGGCGATACTCCAGTGCATTACCAACAGGATATTTTTTGGCATGGTGGATTTTG
 ATGAAGGCTCTGATGTATTTTCTAGATGCTAAACATGAATTCAGCTCCAACCTTTCATCAACTTT
 CCTGCAAAAGGGAAACCCAAACGGGGTGATACATATGAGTTACAGGTGCGGGGTTTTTTCAGC
 TGAGCAGATTGCCCGGTGGATCGCCGACAGAACTGATGTCAATATTAGAGTGATTAGACCCC
 CAAATTATGCTGGTCCCTTATGTTGGGATTGCTTTTGGCTGTTATTGGTGGACTTGTGTAT
 CTTGCAAGAAGTAATATGGAATTTCTCTTTAATAAACTGGATGGGCTTTTGCAGCTTTGTG
 TTTTGTGCTTGCTATGACATCTGGTCAAATGTGGAACCATATAAGAGGACCACCATATGCCC
 ATAAGAATCCCCACACGGGACATGTGAATTATATCCATGGAAGCAGTCAAGCCCAGTTTGTGA
 GCTGAAACACACATTGTTCTTCTGTTAATGGTGGAGTTACCTTAGGAATGGTGTCTTTTATG
 TGAAGCTGCTACCTCTGACATGGATATTGGAAAGCGAAAGATAATGTGTGTGGCTGGTATTG
 GACTTGTGTGATTATTCTTCAGTTGGATGCTCTCTATTTTTTAGATCTAAATATCATGGCTAC
 CCATACAGCTTTCTGATGAGTT**TAA**AAAGGTCCCAGAGATATATAGACACTGGAGTACTGGAA
 ATTGAAAAACGAAAATCGTGTGTGTTTGAAAAGAAGAATGCAACTTGTATATTTTGTATTAC
 CTCTTTTTTTCAAGTGATTTAAATAGTTAATCATTTAAACCAAGAAGATGTGTAGTGCCTTA
 ACAAGCAATCCTCTGTCAAAATCTGAGGTATTTGAAAATAATTATCCTCTTAACCTTCTCTT
 CCCAGTGAACCTTTATGGAACATTTAATTTAGTACAATTAAGTATATTATAAAAAATTGTAAAA
 CTACTACTTTGTTTTAGTTAGAACAAAGCTCAAACTACTTTAGTTAACTTGGTCATCTGAT
 TTTATATTGCCTTATCCAAAGATGGGGAAAGTAAGTCCTGACCAGGTGTTCCACATATGCC
 TGTTACAGATAACTACATTAGGAATTCATTCTTAGCTTCTTCATCTTTGTGTGGATGTGTAT
 ACTTTACGCATCTTTCTTTTGGAGTAGAGAAATTATGTGTGTCATGTGGTCTTCTGAAAATG
 GAACACCATTCTTCAGAGCACACGTCTAGCCCTCAGCAAGACAGTTGTTTCTCCTCCTCCTT
 GCATATTTCTACTGCGCTCCAGCCTGAGTGATAGAGTGAGACTCTGTCTCAAAAAAAGTA
 TCTCTAAATACAGGATTATAATTTCTGCTTGAGTATGGTGTAACTACCTTGTATTTAGAAA
 GATTTTACAGATTCATTCCATCTCCTTAGTTTTCTTTTAAAGGTGACCCATCTGTGATAAAAAATA
 TAGCTTAGTGCTAAAATCAGTGTAACCTTATACATGGCCTAAAATGTTTCTACAAATTAGAGT
 TTGTCACCTTATTCCATTTGTACCTAAGAGAAAAATAGGCTCAGTTAGAAAAGGACTCCCTGG
 CCAGGCGCAGTGACTTACGCCTGTAATCTCAGCACTTTGGGAGGCCAAGGCAGGCAGATCAC
 GAGGTGAGGAGTTCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATAT
 AAAAATTAGCTGGGTGTGGTGGCAGGAGCCTGTAATCCCAGCTACACAGGAGGCTGAGGCAC
 GAGAATCACTTGAACCTCAGGAGATGGAGGTTTTCAGTGAGCCGAGATCACGCCACTGCACTCC
 AGCCTGGCAACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAA

FIGURE 74

MAARWRFWCVSVTMVALLIVCDVPSASAQRKKEMVLSEKVSQLMWETNKRVPVIRMNGDKFR
RLVKAPPRNYSVIVMFTALQLHRQCVVCKQADEEFQILANSWRYSSAFTNRIFAMVDFDEG
SDVFQMLNMNSAPTFINFPAKGKPKRGDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNY
AGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAFALCFVLAMTSGQMWNHIRGPPYAHKN
PHTGHVNYIHGSSQAQFVAETHIVLLFNGGVTLGMVLLCEAATSDMDIGKRKIMCVAGIGLV
VLFFSWMLSIFRSKYHGYPSFLMS

Signal peptide:

amino acids 1-29

Transmembrane domains:

amino acids 183-205, 217-237, 217-287, 301-321

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 75

FIGURE 76

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64885

<subunit 1 of 1, 536 aa, 1 stop

<MW: 61450, pI: 9.17, NX(S/T): 7

MLLLWVSVVAALALAVLAPGAGEQRRRAAKAPNVVLVVSDSFDGRLTFHPGSQVVKLPFINF
MKTRGTSFLNAYTNSPICCPSPRAAMWSGLFTHLTESWNNFKGLDPNYTTWMDVMERHGYRTQ
KFGKLDYTSGHHSISNRVEAWTRDVAFLLRQEGRPMVNLIRNRTKVRVMERDWQNTDKAVNW
LRKEAINYTEPFVIYLG LNLPHYPSPSSGENFGSSTFHTSLYWLEKVSHDAIKIPKWSPLS
EMHPVDYYSYTKNCTGRFTKKEIKNIRAFYYAMCAETDAMLGEIILALHQLDLLQKTIVIIY
SSDHGELAMEHRQFYKMSMYEASAHVPLMMGPGIKAGLQVSNVVS LVDIYPTMLDIAGIPL
PQNLSGYSLPLSSETFKNEHKVKNLHPPWILSEFHGCNVNASTYMLRTNHWKYIAYS DGAS
ILPQLFDLSSDPDEL TNVAVKFPEITYSLDQKLHSIINYPKVSASVHQYNKEQFIKWKQSIG
QNYSNVIANLRWHQDWQKEPRKYENAI DQWLKTHMNPRAV

Important features:**Signal peptide:**

amino acids 1-15

N-glycosylation sites.

amino acids 108-111, 166-169, 193-196, 262-265, 375-378, 413-416,
498-501

Sulfatases proteins:

amino acids 286-315, 359-369, 78-97

FIGURE 77

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAG
AGCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCC**ATG**
GCCTCTCTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACT
GGTTGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTGGTGCCAGCATTGTGACAG
CAGTTGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAG
TGTGACATCTATAGCACCCCTTCTGGGCCTGCCCCGCTGACATCCAGGCTGCCCAGGCCATGAT
GGTGACATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCA
CAGTCTTCTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTC
ATCCTTGGAGGCCTCCTGGGATTCAATTCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGA
CTTCTACTCACCCTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGG
GCATTATTTCTTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCCTGCTCATCC
CAGAGAAATCGCTCCAACCTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTC
TCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCTACAGCCTGACAGGGT
ATGTG**TGA**AGAACCAGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAG
CACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGTGTGAGAAGGTGCTGCTGAGG
ATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATG
CAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTC
CCCTGCCCTAAGTCCCCAACCCTCAACTTGAAACCCCATTCCTTAAGCCAGGACTCAGAGG
ATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCACTG
ACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGG
GGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCC
TCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCC
AGACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAG
GATGCAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 78

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64886

><subunit 1 of 1, 230 aa, 1 stop

><MW: 24549, pI: 8.56, NX(S/T): 1

MASLGLQLVGYILGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGIT
QCDIYSTLLGLPADIQAAQAMMVTSSAISLACIISVVGMRCTVFCQESRAKDRVAVAGGVF
FILGGLLGFIPIVAVNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCS
SQNRNSNYDAYQAQPLATRSSPRPGQPPKVKSEFNLSYSLTGYV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 79

GCACTGCTGCTGTCCCATCAGCTGCTCTGAAGCTCCATGGTGCCCAGAATCTTCGCTCCTGC
TTATGTGTCACTCTGTCTCCTCCTCTTGTGTCCAAGGGAAGTCATCGCTCCCGCTGGCTCAG
AACCATGGCTGTGCCAGCCGGCACCCAGGTGTGGAGACAAGATCTACAACCCCTTGGAGCAG
TGCTGTTACAATGACGCCATCGTGTCCCTGAGCGAGACCCGCCAATGTGGTCCCCCCTGCAC
CTTCTGGCCCTGCTTTGAGCTCTGCTGTCTTGATTCCTTTGGCCTCACAAACGATTTTGTG
TGAAGCTGAAGGTTCAAGGTGTGAATTCCCAGTGCCACTCATCTCCCATCTCCAGTAAATGT
GAAAGCAGAAGACGTTTTCCCTGAAGACATAGAAAGAAAATCAACTTTACTAAGGCATC
TCAGAAACATAGGCTAAGGTAATATGTGTACCAGTAGAGAAGCCTGAGGAATTTACAAAATG
ATGCAGCTCCAAGCCATTGTATGGCCCATGTGGGAGACTGATGGGACATGGAGAATGACAGT
AGATTATCAGGAAATAAATAAAGTGGTTTTTCCAATGTACACACCTGTAAAA

FIGURE 80

MVPRI FAPAYVSVCLLLLCPREVIAPAGSEPWLCQPAPRCGDKIYNPLEQCCYNDAIVSLSE
TRQCGPPCTFWPCFELCCLDSFGLTNDVFVVKLVQGVNSQCHSSPISSKCESRRRFP

Signal peptide:

amino acids 1-25

80/249

FIGURE 81

CTCCACTGCAACCACCCAGAGCCATGGCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCATT
TTCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTTACCT
GATGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCCTGCAGCACTGTT
GCTATGATGATGCCGTCGTGCCCTTGGCCAGGACCCAGACGTGTGGAAACTGCACCTTCAGA
GTCTGCTTTGAGCAGTGCTGCCCCCTGGACCTTCATGGTGAAGCTGATAAACCAGAACTGCCA
CTCAGCCCCGGACCTCGGATGACAGGCTTTGTGCGAGTGTGAGCTAATGGAACATCAGGGGAA
CGATGACTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCTGGTGTTACCTGAGA
TCTGGGATGCTGAGTGGCTGTTTGGGGGCCAGAGAAACACACACTCAACTGCCCACCTTCATT
CTGTGACCTGTCTGAGGCCACCCTGCAGCTGCCCTGAGGAGGCCACAGGTCCCCTTCTAG
AATTCTGGACAGCATGAGATGCGTGTGCTGATGGGGGCCAGGGACTCTGAACCCTCCTGAT
GACCCCTATGGCCAACATCAACCCGGCACCACCCCAAGGCTGGCTGGGGAACCCTTCACCCT
TCTGTGAGATTTTCCATCATCTCAAGTTCTCTTCTATCCAGGAGCAAAGCACAGGATCATAA
TAAATTTATGTACTTTATAAATGAAAA

FIGURE 82

MAPRGCIVAVFAIFCISRLLC SHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAVVP
LARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCDSARTSDDRLCRSVS

Signal peptide:

amino acids 1-24

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 83

GGGGGCGGGTGCCTGGAGCACGGCGCTGGGGCCGCCCCGAGCGCTCACTCGCTCGCACTCAG
 TCGCGGGAGGCTTCCCCGCGCCGGCCGCGTCCCGCCCGCTCCCGGGCACCAGAAGTTCCTCT
 GCGCGTCCGACGGCGAC**ATG**GGCGTCCCCACGGCCCTGGAGGCCGGCAGCTGGCGCTGGGGA
 TCCCTGCTCTTCGCTCTCTTCCTGGCTGCGTCCCTAGGTCCGGTGGCAGCCTTCAAGGTCGC
 CACGCCGTATTCCCTGTATGTCTGTCCCGAGGGGCAGAACGTCACCCTCACCTGCAGGCTCT
 TGGGCCCTGTGGACAAAGGGCACGATGTGACCTTCTACAAGACGTGGTACCGCAGCTCGAGG
 GCGAGGTGCAGACCTGCTCAGAGCGCCGGCCCATCCGCAACCTCACGTTCCAGGACCTTCA
 CCTGCACCATGGAGGCCACCAGGCTGCCAACACCAGCCACGACCTGGCTCAGCGCCACGGGC
 TGGAGTCGGCCTCCGACCACCATGGCAACTTCTCCATCACCATGCGCAACCTGACCCTGCTG
 GATAGCGGCCTCTACTGCTGCCTGGTGGTGGAGATCAGGCACCACCCTCGGAGCACAGGGT
 CCATGGTGCCATGGAGCTGCAGGTGCAGACAGGCCAAAGATGCACCATCCAAGTGTGTGGTGT
 ACCCATCCTCCTCCCAGGATAGTGAAAACATCACGGCTGCAGCCCTGGCTACGGGTGCCTGC
 ATCGTAGGAATCCTCTGCCTCCCCCTCATCCTGCTCCTGGTCTACAAGCAAAGGCAGGCAGC
 CTCCAACCGCCGTGCCAGGAGCTGGTGC GGATGGACAGCAACATTCAAGGGATTGAAAACC
 CCGGCTTTGAAGCCTCACACCTGCCCAGGGGATAACCGAGGCCAAAGTCAGGCACCCCTG
 TCCTATGTGGCCCAGCGGCAGCCTTCTGAGTCTGGGCGGCATCTGCTTTCGGAGCCCAGCAC
 CCCCCTGTCTCCTCCAGGCCCCGGAGACGTCTTCTTCCCATCCCTGGACCCTGTCCCTGACT
 CTCCAAACTTTGAGGTCATC**TAG**CCCAGCTGGGGGACAGTGGGCTGTTGTGGCTGGGTCTGG
 GGCAGGTGCATTTGAGCCAGGGCTGGCTCTGTGAGTGGCCTCCTTGGCCTCGGCCCTGGTTC
 CCTCCCTCCTGCTCTGGGCTCAGATACTGTGACATCCCAGAAGCCCAGCCCCTCAACCCCTC
 TGGATGCTACATGGGGATGCTGGACGGCTCAGCCCCTGTTCCAAGGATTTTGGGGTGTGAG
 ATTCTCCCCCTAGAGACCTGAAATTCACCAGCTACAGATGCCAAATGACTTACATCTTAAGAA
 GTCTCAGAACGTCCAGCCCTTCAGCAGCTCTCGTTCTGAGACATGAGCCTTGGGATGTGGCA
 GCATCAGTGGGACAAGATGGACACTGGGCCACCCTCCCAGGCACCAGACACAGGGCACGGTG
 GAGAGACTTCTCCCCCGTGGCCGCCTTGGCTCCCCCGTTTTGCCCGAGGCTGCTCTTCTGTC
 AGACTTCCTCTTTGTACCACAGTGGCTCTGGGGCCAGGCCTGCCTGCCCACTGGCCATCGCC
 ACCTTCCCCAGCTGCCTCCTACCAGCAGTTTCTCTGAAGATCTGTCAACAGGTTAAGTCAAT
 CTGGGGCTTCCACTGCCTGCATTCCAGTCCCCAGAGCTTGGTGGTCCCGAAACGGGAAGTAC
 ATATTGGGGCATGGTGGCCTCCGTGAGCAAATGGTGTCTTGGGCAATCTGAGGCCAGGACAG
 ATGTTGCCCCACCCACTGGAGATGGTGTCTGAGGGAGGTGGGTGGGGCCTTCTGGGAAGGTGA
 GTGGAGAGGGGCACCTGCCCCCGCCCTCCCCATCCCCTACTCCCCTGCTCAGCGCGGGCC
 ATTGCAAGGGTGCCACACAATGTCTTGTCCACCCTGGGACACTTCTGAGTATGAAGCGGGAT
 GCTATTA AAAACTACATGGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

FIGURE 84

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64897

><subunit 1 of 1, 311 aa, 1 stop

><MW: 33908, pI: 6.87, NX(S/T): 6

MGVPTALEAGSWRWGSLLFALFLAASLGPVAAFKVATPYSLYVCPEGQNVTLTCRLLGPVDK
GHDVTFYKTWYRSSRGVQTCSERRPIRNLTFQDLHLHHGGHQAANTSHDLAQRHGLESASD
HHGNFSITMRNLTLDSGLYCCLVVEIRHHHSEHRVHGAMELQVQTGKDAPSNVCVYPSSSQ
DSENITAAALATGACIVGILCLPLILLVYKQRQAASNRRAQELVRMDSNIQGIENPGFEAS
PPAQGIPEAKVRHPLSYVAQRQPSESGRHLLSEPSTPLSPPGPDVFFPSLDPVPDSPNFEVI

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 190-216

FIGURE 85

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCT
 TTCCCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCTGCCCTCC
 TTTCTGCCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGG
 GGTCTGTGGGTTGATCTGTGGCCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCCTCCCGA
 CTCCGCTCCCGGACCAGCGGCCTGACCCTGGGGAAAGG**ATG**GTTCCCGAGGTGAGGGTCCTC
 TCCTCCTTGCTGGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCCGCC
 AGACATGTTCTGCCTTTTCCATGGGAAGAGATACTCCCCCGGCGAGAGCTGGCACCCCTACT
 TGGAGCCACAAGGCCTGATGTACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGT
 TGTACCGCCTCCACTGTCCGCCTGTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATG
 CTGTCCCAAGTGTGTGGAACCTCACACTCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCC
 AGCACAACGGGACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTTCCCTCC
 CGCCTGCCCAACCAGTGTGTCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCCTCAC
 AACCTGCCCCGAACCAGGCTGCCCAGCACCCCTCCCACTGCCAGACTCCTGCTGCCAAGCCT
 GCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTG
 AGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAGAGAGGGCCCGGGCACCCAGC
 CCCCCTGGCCTCAGCGCCCCCTCTGAGCTTCATCCCTCGCCACTTCAGACCCAAGGGAGCAG
 GCAGCACAACCTGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGGCGGG
 AAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGCCCCCTTGCCCTG
 CATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGT
 ACCCCTGCCGTCACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTTGCCCAGAGGACAAA
 GCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTCTT
 CGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGG
 CCTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAG
 AGAGGTGAAGTACCTGGCCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGA
 AAGTCAGGAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCAC
 GAAGGTCACTGGAACGTCTTCCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGA
 CAAAGTGACCAAGACATAACAAAGACC**TAA**CAGTTGCAGATATGAGCTGTATAATTGTTGTT
 ATTATATATTAATAAATAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 86

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64902

><subunit 1 of 1, 451 aa, 1 stop

><MW: 49675, pI: 7.15, NX(S/T): 1

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRCT
CSEGAHVSCYRLHCPPVHCPQPVTEPQQCCPKC VEPHTPSGLRAPPKSCQHNGTMYQHGEIF
SAHELFPSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEED
SVQSLHGVRHPQDPCSSDAGRKRGP GTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKH
KKACVHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPCRHPK VAGKC
CKICPEDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKL V
KDEETEAQRGEVPGPRPHSQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSDPG
AEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 87

CTAGCCTGCGCCAAGGGGTAGTGAGACCGCGCGGCAACAGCTTGCGGCTGCGGGGAGCTCCC
GTGGGCGCTCCGCTGGCTGTGCAGGCGGCC**ATG**GATTCCCTTGCGGAAAATGCTGATCTCAGT
CGCAATGCTGGGCGCAGGGGCTGGCGTGGGCTACGCGCTCCTCGTTATCGTGACCCCGGGAG
AGCGGCGGAAGCAGGAAATGCTAAAGGAGATGCCACTGCAGGACCCAAGGAGCAGGGAGGAG
GCGGCCAGGACCCAGCAGCTATTGCTGGCCACTCTGCAGGAGGCAGCGACCACGCAGGAGAA
CGTGGCCTGGAGGAAGAACTGGATGGTTGGCGGCGAAGGCGGCGCCAGCGGGAGGTCACCG**T**
GAGACCGGACTTGCCTCCGTGGGCGCCGGACCTTGGCTTGGGCGCAGGAATCCGAGGCAGCC
TTTCTCCTTCGTGGGCCCAGCGGAGAGTCCGGACCGAGATAACCATGCCAGGACTCTCCGGGG
TCCTGTGAGCTGCCGTGCGGTGAGCACGTTTCCCCCAAACCCTGGACTGACTGCTTTAAGGT
CCGCAAGGCGGGCCAGGGCCGAGACGCGAGTCGGATGTGGTGAAGTGAAGAACCAATAAAA
TCATGTTCTCCAAA
AAAAAAAAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 88

MDSLRLKMLISVAMLGAGAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAARTQQLLLA
TLQEAATTQENVAWRKNWMVGEGGASGRSP

Signal peptide:

amino acids 1-18

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 89

[illegible]

FIGURE 90

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64905

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46038, pI: 6.50, NX(S/T): 2

MGPSTPLLILFLLSWSGPLQGQQHHLVEYMERRLAALAEERLAQCQDQSSRHAAELRDFKNKM
LPILLEVAEKEREALRTEADTISGRVDRLEREVDYLETQNPALPCVEFDEKVTGGPGTKGKGR
RNEKYDMVTDGCGYTISQVRSMKILKRFGGPAGLWTKDPLGQTEKIYVLDGTQNDTAFVFPRL
RDFTLAMAARKASRVRVPFPWVG TGQLVYGGFLYFARRPPGRPGGGEMENTLQLIKFHLAN
RTVVDSSVFPAEGLIPPYGLTADTYIDLVADEEGLWAVYATREDDRHLCLAKLDPQTLDTQ
QWDTPCPRENAEAAAFVICGTLYVVYNTRPASRARIQCSFDASGTLTPERAALPYFPRRYGAH
ASLRYNPRERQLYAWDDGYQIVYKLEMRKKEEEV

Important features:**Signal peptide:**

amino acids 1-21

N-glycosylation sites.

amino acids 177-180, 248-251

FIGURE 91

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAG
 AGCCTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTAATCCAT
 CCGTCACCTCTCCTGTCATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCC**ATGG**
 CTCTCATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTT
 GGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAGGACGCAGCATTCTCCTGTTTCCTGTC
 TCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGGCCAGTTCTCTAGCGTGG
 TCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAAGGCAGG
 ACAAACCTGGTGAAGGATTCTATTGCGGAGGGGGCGCATCTCTCTGAGGCTGGAAAACATTAC
 TGTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCA
 TCTGGGAGCTACAGGTGTCAGCACTGGGCTCAGTTCCTCTCATTTCCATCACGGGATATGTT
 GATAGAGACATCCAGCTACTCTGTGAGTCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTG
 GAAAGGTCCACAAGGACAGGATTTGTCCACAGACTCCAGGACAAACAGAGACATGCATGGCC
 TGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAACGCCGGGAGCATATCCTGTTCCATG
 CGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAGGAGATACCTTTTTTCGA
 GCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTATTTTTTG
 GCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCAGTGGAAAATCCAGGCGGAACTGGAC
 TGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGAC
 TCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTGAACCCATA
 GAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCT
 TCTCAGAGTTTCCAAGCAGGGAAACATTACTGGGAGGTGGACGGAGGACACAATAAAAGGTG
 GCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCCCG
 ATCATGGGTACTGGGTCCTCAGACTGAATGGAGAACATTTGTATTTACATTAAATCCCCGT
 TTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCCTGGACTATGAGTG
 TGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGT
 TTGAAGGCTTATTGAGGCCCTACATTGAGTATCCGTCCTATAATGAGCAAAATGGAACTCCC
 ATAGTCATCTGCCCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGC
 AATCCCAGAGACAAGCAACAGTGAGTCCTCCTCACAGGCAACCACGCCCTTCCTCCCCAGGG
 GTGAAATG**TAG**GATGAATCACATCCCACATTCTTCTTTAGGGATATTAAGGTCTCTCTCCCA
 GATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCAGATGAAGGGGGACTGGCCTGTCC
 ACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGGCTGACATTACATTT
 AGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAGAACCG
 TCAGGAATTCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGC
 TTAGATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTA
 AAAAAA

[illegible][illegible]

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

[illegible][illegible][illegible]

FIGURE 93

GCGATGGTGC GCCCGGTGGCGGTGGCGGCGGCGGTTGCGGAGGCTTCCTTGGTCCGATTGCA
 ACGAGGAGAAGATGACTGACCAACCGACTGGCTGAATGAATGAATGGCGGAGCCGAGCGCGC
CATGAGGAGCCTGCCGAGCCTGGGCGGCCTCGCCCTGTTGTGCTGCGCCGCCGCCGCCGCCG
 CCGTCGCCTCAGCCGCCTCGGCGGGGAATGTCACCGGTGGCGGCGGGGCCGCGGGGCAGGTG
 GACGCGTCGCCGGGCCCCGGGTGCGGGGCGAGCCCAGCCACCCCTTCCCTAGGGCGACGGC
 TCCCACGGCCCCAGGCCCCGAGGACCGGGCCCCCGCGCGCCACCGTCCACCGACCCCTGGCTG
 CGACTTCTCCAGCCAGTCCCCGGAGACCACCCCTCTTTGGGCGACTGCTGGACCCTCTTCC
 ACCACCTTTTCAGGCGCCGCTCGGCCCTCGCCGACCACCCCTCCGGCGGCGGAACGCACTTC
 GACCACCTCTCAGGCGCCGACCAGACCCGCGCCGACCACCCCTTTCGACGACCACTGGCCCCG
 CGCCGACCACCCCTGTAGCGACCACCGTACCGGCGCCACGACTCCCCGGACCCCGACCCCC
 GATCTCCCCAGCAGCAGCAACAGCAGCGTCTCCCCACCCACCTGCCACCGAGGCCCCCTC
 TTCGCTCCTCCAGAGTATGTATGTAACCTGCTCTGTGGTTGGAAGCCTGAATGTGAATCGCT
 GCAACCAGACCACAGGGCAGTGTGAGTGTGCGGCCAGGTTATCAGGGGCTTCACTGTGAAACC
 TGCAAAGAGGGCTTTTACCTAAATTACACTTCTGGGCTCTGTGAGCCATGTGACTGTAGTCC
 ACATGGAGCTCTCAGCATACCGTGCAACAGG**TAAG**CAACAGAGGGTGGAAGTGAAGTTTATT
 TTATTTTAGCAAGGGAAAAAAAAGGCTGCTACTCTCAAGGACCATACTGGTTTAAACAAAG
 GAGGATGAGGGTCATAGATTTACAAAATATTTTATATACTTTTATTCTTACTTTTATATGT
 TATATTTAATGTCAGGATTTAAAAACATCTAATTTACTGATTTAGTTCTTCAAAAGCACTAG
 AGTCGCCAATTTTCTCTGGGATAATTTCTGTAAATTTTCATGGGAAAAAATTATTGAAGAAT
 AAATCTGCTTTCTGGAAGGGCTTTCAGGCATGAAACCTGCTAGGAGGTTTAGAAATGTTCTT
 ATGTTTATTAATATAACCATTGGAGTTTGAGGAAATTTGTTGTTTGGTTTATTTTTCTCTCTA
 ATCAAAATTCTACATTTGTTTCTTTGGACATCTAAAGCTTAACCTGGGGGTACCCTAATTTA
 TTTAACTAGTGGTAAGTAGACTGGTTTTACTCTATTTACCAGTACATTTTTTGAGACCAAAG
 TAGATTAAGCAGGAATTATCTTTAAACTATTATGTTATTTGGAGGTAATTTAATCTAGTGGA
 ATAATGTACTGTTATCTAAGCATTTGCCTTGTAAGTACTGCAAGTAATTTATCTTTGACCT
 TATGTGAGGCACTTGGCTTTTTGTGGACCCCAAGTCAAAAAACTGAAGAGACAGTATTAAAT
 AATGAAAAAAATAATGACAGGTTATACTCAGTGTAACCTGGGTATAACCCAAGATCTGCTGC
 CACTTACGAGCTGTGTTCCCTTGGGCAAGTAATTTCCCTTCACTGAGCTTGTTTCTTCTCAAG
 GTTGTGTTGTAAGATTAAATGAGTTGATATATATAAAATGCCTAGCACATGTCACTCAATAAA
 TTCTGGTTTGTTTTAAATTTCAAAGGAATATTATGGACTGAAATGAGAGAACATGTTTAAAGA
 ACTTTTAGCTCCTTGACAAAGAGTGCTTTATACTTTAGCACTAAATATTTTAAATGCTTTA
 TAAATGATATTATACTGTTATGGAATATTGTATCATATTGTAGTTTATTAATAAATGTAGAAG
 AGGCTGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTTTGGGAGGCCAAGGCGGGTGGAT
 CACTTGAGGCCAGGAGTTCTAGATGAGCCTGGCCAGCACAGTGAAACCCCGTCTCTACTAAA
 AATACAAACAAATTAGCTGGGCGTGGTGGCACACACCTGTAGTCCCAGCTACTCGGGAGGCT
 GAGGCAGGAGAATCGGTTGAACCCGGGAGGTGGAGGTGCAGTGAGCTGAGATCGCGCCACT
 GCACTCCAGCCTGGTGAGAGAGGGAGACTCTGTCTTAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 94

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64952

><subunit 1 of 1, 258 aa, 1 stop

><MW: 25716, pI: 8.13, NX(S/T): 5

M R S L P S L G G L A L L C C A A A A A V A S A A S A G N V T G G G G A A G Q V D A S P G P G L R G E P S H P F P R A T A
P T A Q A P R T G P P R A T V H R P L A A T S P A Q S P E T T P L W A T A G P S S T T F Q A P L G P S P T T P P A A E R T S
T T S Q A P T R P A P T T L S T T T G P A P T T P V A T T V P A P T T P R T P T P D L P S S S N S S V L P T P P A T E A P S
S P P E Y V C N C S V V G S L N V N R C N Q T T G Q C E C R P G Y Q G L H C E T C K E G F Y L N Y T S G L C Q P C D C S P
H G A L S I P C N R

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 30-33, 172-175, 195-198, 208-211, 235-238

EGF-like domain cysteine pattern signature.

amino acids 214-226.

FIGURE 95

TGCGGCGCAGTGTAGACCTGGGAGGATGGGCGGCCTGCTGCTGGCTGCTTTTCTGGCTTTGG
TCTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCTTGGG
CCCTGGTACGTGCTTGCGGTGGCCTCCCGGGAAAAGGGCTTTGCCATGGAGAAGGACATGAA
GAACGTCGTGGGGGTGGTGGTGACCCTCACTCCAGAAAACAACCTGCGGACGCTGTCCTCTC
AGCACGGGCTGGGAGGGTGTGACCAGAGTGTTCATGGACCTGATAAAGCGAAACTCCGGATGG
GTGTTTGAGAATCCCTCAATAGGCGTGCTGGAGCTCTGGGTGCTGGCCACCAACTTCAGAGA
CTATGCCATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTTCAACACCGTGGAGCTGT
ACAGTCTGACGGAGACAGCCAGCCAGGAGGCCATGGGGCTCTTCACCAAGTGGAGCAGGAGC
CTGGGCTTCCTGTCACAGTAGCAGGCCCAGCTGCAGAAGGACCTCACCTGTGCTCACAAGAT
CCTTCTGTGAGTGCTGCGTCCCCAGTAGGGATGGCGCCACAGGGTCCTGTGACCTCGGCCA
GTGTCCACCCACCTCGCTCAGCGGCTCCCGGGGCCAGCACCAGCTCAGAATAAAGCGATTC
CACAGCA

FIGURE 96

MGGLLLAAFLALVSVPRQA VWLGRLDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVT
LTPENNLRTLSSQHGLGGCDQSVMDLIKRN SGWVFENPSIGVLELWVLATNFRDYAIIFTQL
EFGDEPFNTVELYSLTETASQEAMGLFTKWSRSLGFLSQ

Signal peptide:

amino acids 1-20

Downloaded from www.sciencedirect.com

FIGURE 97

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCCAGAC**ATG**CTGCTGCTGCTGCTGCC
CCTGCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTT
CCGTGACGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCAT
GGCTGGATTTACCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGA
CCAGGATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACC
GATTCCACCTCCTTGGGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGA
AGAAGTGATGCGGGGAGATACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAA
ACATCACCGGCTCTCTGTGAATGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAG
GCACCCTGGAGTCCGGCTGCCCCCAGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAG
GGGACACCCCCTATGATCTCCTGGATAGGGACCTCCGTGTCCCCCTGGACCCCTCCACCAC
CCGCTCCTCGGTGCTCACCCCTCATCCACAGCCCCAGGACCATGGCACCAGCCTCACCTGTC
AGGTGACCTTCCCTGGGGCCAGCGTGACCACGAACAAGACCGTCCATCTCAACGTGTCCTAC
CCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGAGACGGCACAGTATCCACAGTCTTGGG
AAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCTGCGCCTGGTCTGTGCAGTTGATG
CAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGAGAGGCCTGACCCTGTGCCCC
TCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCACCTGAGGGATGCAGCTGA
ATTCACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCTGAACGTCTCCCTGC
AGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTGGGGGAGCTGGAGCCACAGCCCTG
GTCTTCCTGTCCTTCTGCGTCATCTTCGTTGTAGTGAGGTCCTGCAGGAAGAAATCGGCAAG
GCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTCAGCCT
CTCAGGGGCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAGCT
TCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGT
GAAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCC
ACAGAT**TGA**GAAACTGCAGAGACTCACCCCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGA
GAAGTCAGAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACT
ATGAATTATGTGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAAT
CCACACTGTGCCCTCCCTTTTATTTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 98

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSPSHGWIYPGPVVHGYWF
REGANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKG
SIKWNYKHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVS
PLDPSTTRSSVLTLIPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPQNLMTVFQGDG
TVSTVLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWV
HLRDAAEFTCRAQNPLGSQQVYLNVSLSQKATSGVTQGVVGGAGATALVFLSFCVIFVVVRS
CRKKSARPAAGVGDTGIEDANAVRGSSASQGPLEPWAEDSPPDQPPPASARSSVGEDELQYA
SLSFQMVKPWDSRGQEATDTEYSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 99

GACGCCCAGTGACCTGCCGAGGTCGGCAGCACAGAGCTCTGGAGATGAAAGACCCTGTTTCCTG
GGTGTACGCTCGGCCTGGCCGCTGCCCTGTCCTTCACCCTGGAGGAGGAGGATATCACAGG
GACCTGGTACGTGAAGGCCATGGTGGTCGATAAGGACTTTCCGGAGGACAGGAGGCCAGGA
AGGTGTCCCCAGTGAAGGTGACAGCCCTGGGCGGTGGGAAGTTGGAAGCCACGTTTCACCTTC
ATGAGGGAGGATCGGTGCATCCAGAAGAAAATCCTGATGCGGAAGACGGAGGAGCCTGGCAA
ATACAGCGCCTATGGGGGCAGGAAGCTCATGTACCTGCAGGAGCTGCCCAGGAGGGACCACT
ACATCTTTTACTGCAAAGACCAGCACCATGGGGGCCTGCTCCACATGGGAAAGCTTGTGGGT
AGGAATTCTGATACCAACCGGGAGGCCCTGGAAGAATTAAAGAAATTGGTGCAGCGCAAGGG
ACTCTCGGAGGAGGACATTTTACGCCCCCTGCAGACGGGAAGCTGCGTTCCCGAACACTAGG
CAGCCCCCGGGTCTGCACCTCCAGAGCCCACCCTACCACCAGACACAGAGCCCGGACCACCT
GGACCTACCCTCCAGCCATGACCCTTCCCTGCTCCCACCCACCTGACTCCAAATAAAGTCCT
TTTCCCCCAA

FIGURE 100

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65404

<subunit 1 of 1, 170 aa, 1 stop

<MW: 19457, pI: 9.10, NX(S/T): 0

MKTLFLGVTLGLAAALSFTLEEDITGTWYVKAMVVDKDFPEDRRPRKVSPVKVTALGGGKL

EATFTFMREDRCIQKKILMRKTEEPGKYSAYGGRKLMYLQELPRRDHYIFYCKDQHHGGLLH

MGKLVGRNSDTNREALEEFKKLVQRKGLSEEDIFTPLQTGSCVPEH

Important features:

Signal peptide:

amino acids 1-17

100/249

FIGURE 101

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTCAC
AGCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCTCCAGGCC**ATG**AGGATTCTGCAGTTAA
TCCTGCTTGCTCTGGCAACAGGGCTTGTAGGGGGAGAGACCAGGATCATCAAGGGGTTCGAG
TGCAAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTGTTGAGAAGACGCGGGCTACTCTGTGG
GGCGACGCTCATCGCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAGCCCCGCTACA
TAGTTCACCTGGGGCAGCACAACCTCCAGAAGGAGGAGGGCTGTGAGCAGACCCGGACAGCC
ACTGAGTCCTTCCCCCACCCCGGCTTCAACAACAGCCTCCCCAACAAGACCACCGCAATGA
CATCATGCTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCC
TCTCCTCACGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCAGCACGTCC
AGCCCCCAGTTACGCCTGCCTCACACCTTGCGATGCGCCAACATCACCATCATTGAGCACCA
GAAGTGTGAGAACGCCTACCCCGGCAACATCACAGACACCATGGTGTGTGCCAGCGTGCAGG
AAGGGGGCAAGGACTCCTGCCAGGGTGACTCCGGGGGGCCCTCTGGTCTGTAACCAGTCTCTT
CAAGGCATTATCTCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACAC
GAAAGTCTGCAAATATGTGGACTGGATCCAGGAGACGATGAAGAACAAT**TAG**ACTGGACCCA
CCCACCACAGCCCATCACCTCCATTTCCACTTGGTGTTTGGTTCCTGTTCACTCTGTTAAT
AAGAAACCCTAAGCCAAGACCCTCTACGAACATTCTTTGGGCCTCCTGGACTACAGGAGATG
CTGTCACTTAATAATCAACCTGGGGTTCGAAATCAGTGAGACCTGGATTCAAATTCTGCCTT
GAAATATTGTGACTCTGGGAATGACAACACCTGGTTTGTCTCTGTTGTATCCCCAGCCCCA
AAGACAGCTCCTGGCCATATATCAAGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 102

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65405

<subunit 1 of 1, 250 aa, 1 stop

<MW: 27466, pI: 8.87, NX(S/T): 4

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAH
 CLKPRYIVHLGQHNLOKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSITW
 AVRPLTLSSRCVTAGTSC LISGWGSTSSPQLRLPHTLRCANITIIEHQKCENAYPGNITDTM
 VCASVQEGGKDSCQGDSSGGLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYVDWIQETMKNN

Important features:**Signal peptide:**

amino acids 1-18

Serine proteases, trypsin family, histidine active site.

amino acids 58-63

N-glycosylation sites.

amino acids 99-102, 165-168, 181-184, 210-213

Glycosaminoglycan attachment site.

amino acids 145-148

Kringle domain proteins.

amino acids 197-209, 47-64

Serine proteases, trypsin family, histidine protein

amino acids 199-209, 47-63, 220-243

Apple domain proteins

amino acids 222-249, 189-222

FIGURE 103

GAGCAGTGTTCTGCTGGAGCCG**ATG**CCAAAAACCATGCATTTCTTATTCAGATTCATTGTTT
TCTTTTATCTGTGGGGCCTTTTTACTGCTCAGAGACAAAAGAAAGAGGAGAGCACCGAAGAA
GTGAAAATAGAAGTTTTGCATCGTCCAGAAAAGTCTCTAAGACAAGCAAGAAGGGAGACCT
ACTAAATGCCCATTATGACGGCTACCTGGCTAAAGACGGCTCGAAATTCTACTGCAGCCGGA
CACAAAATGAAGGCCACCCCAAATGGTTTGTCTTGGTGTGGGCAAGTCATAAAAGGCCTA
GACATTGCTATGACAGATATGTGCCCTGGAGAAAAGCGAAAAGTAGTTATACCCCTTCATT
TGCATACGGAAAGGAAGGCTATGCAGAAGGCAAGATTCACCGGATGCTACATTGATTTTTG
AGATTGAACTTTATGCTGTGACCAAAGGACCACGGAGCATTGAGACATTTAAACAAATAGAC
ATGGACAATGACAGGCAGCTCTCTAAAGCCGAGATAAACCTCTACTTGCAAAGGGAATTTGA
AAAAGATGAGAAGCCACGTGACAAGTCATATCAGGATGCAGTTTTAGAAAGATATTTTTAAGA
AGAATGACCATGATGGTGATGGCTTCATTTCTCCCAAGGAATACAATGTATACCAACACGAT
GAACTA**TAG**CATATTTGTATTTCTACTTTTTTTTTTTTAGCTATTTACTGTACTTTATGTATA
AAACAAAGTCACTTTTCTCCAAGTTGTATTTGCTATTTTCCCCTATGAGAAGATATTTTGA
TCTCCCAATACATTGATTTTGGTATAATAAATGTGAGGCTGTTTTGCAAACCTTAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 104

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65406

<subunit 1 of 1, 222 aa, 1 stop

<MW: 25794, pI: 6.24, NX(S/T): 1

MPKTMHFLFRFIVFFYLLWGLFTAQRQKKEESTEEVKIEVLHRPENCSKTSKKGDLLNAHYDG
YLAKDGSKFYCSRTQNEGHPKWFVLGVGVQVIKGLDIAMTDMCPGEKRKVVIPPSFAYGKEGY
AEGKIPPDATLIFEIELYAVTKGPRSIETFKQIDMDNDRQLSKAEINLYLQREFEKDEKPRD
KSYQDAVLEDIFKKNDHGDGDFISPKEYNVYQHDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 219-222

N-glycosylation site.

amino acids 45-48

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 87-223, 129-142

EF-hand calcium-binding domain proteins

amino acids 202-214, 195-214

[illegible][illegible]

FIGURE 106

MQGPLLLPGLCFLLSLFGAVTQKTKTSCAKCPPNASCVNNTHCTCNHGYTSGSGQKLFTFPL
ETCNARHGG SRL

Signal peptide:

amino acids 1-18

106/249

FIGURE 107

CAAGCAGGTCATCCCCTTGGTGACCTTCAAAGAGAAGCAGAGAGGGCAGAGGTGGGGGGGCAC
AGGGAAAGGGTGACCTCTGAGATTCCCCTTTTCCCCCAGACTTTGGAAGTGACCCACCATGG
GGCTCAGCATCTTTTTGCTCCTGTGTGTTCTTGGGCTCAGCCAGGCAGCCACACCCGAAGATT
TTCAATGGCACTGAGTGTGGGCGTAACTCACAGCCGTGGCAGGTGGGGCTGTTTGAGGGGCAC
CAGCCTGCGCTGCGGGGGTGTCTTATTGACCACAGGTGGGTCTCACAGCGGCTCACTGCA
GCGGCAGCAGGTACTGGGTGCGCCTGGGGGAACACAGCCTCAGCCAGCTCGACTGGACCGAG
CAGATCCGGCACAGCGGCTTCTCTGTGACCCATCCCGGCTACCTGGGAGCCTCGACGAGCCA
CGAGCACGACCTCCGGCTGCTGCGGCTGCGCCTGCCCCTCCGCGTAACCAGCAGCGTTCAAC
CCCTGCCCCCTGCCCAATGACTGTGCAACCGCTGGCACCAGTGCCACGTCTCAGGCTGGGGC
ATCACCAACCACCCACGGAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGT
CTCCCATGCCACCTGCCATGGTGTGTATCCCGGGAGAATCACGAGCAACATGGTGTGTGCAG
GCGGCGTCCCGGGGCAGGATGCCTGCCAGGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGGA
GTCCTTCAAGGTCTGGTGTCTTGGGGGTCTGTGGGGGCCCTGTGGACAAGATGGCATCCCTGG
AGTCTACACCTATATTTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAACTGAC
CTGTTTCCTCCACCTCCACCCCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACC
AATATCTCCTCCATCACTTCCCCTAGCTCCACTCTTGTGGCCTGGGAACTTCTTGGAACCT
TAACTCCTGCCAGCCCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGA
ATAAATATAAATGAAGGAGGGGCAAAAAAAAAAAAAA

FIGURE 108

MGLSIFLLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLTAAH
CSGSRYWVRLGEHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLLRLRLPVRVTSSV
QPLPLPNDCATAGTECHVSGWGITNHPRNPFDPDLLQCLNLSIVSHATCHGVYPGRITSNMVC
AGGVPGQDACQGDSGGPLVCGGVLQGLVSWGSGVPCGQDGIPGVYTYICKYVDWIRMIMRNN

Signal peptide:

amino acids 1-17

108/249

FIGURE 109

GCGGCCACACGCAGCTAGCCGGAGCCCGGACCAGGCGCCTGTGCCTCCTCCTCGTCCCTCGC
CGCGTCCGCGAAGCCTGGAGCCGGCGGGAGCCCCGCGCTCGCC**ATG**TCGGGCGAGCTCAGCA
ACAGGTTCCAAGGAGGGAAGGCGTTCGGCTTGCTCAAAGCCCGGCAGGAGAGGAGGCTGGCC
GAGATCAACCGGGAGTTTCTGTGTGACCAGAAGTACAGTGATGAAGAGAACCTTCCAGAAAA
GCTCACAGCCTTCAAAGAGAAGTACATGGAGTTTGACCTGAACAATGAAGGCGAGATTGACC
TGATGTCTTTAAAGAGGATGATGGAGAAGCTTGGTGTCCCCAAGACCCACCTGGAGATGAAG
AAGATGATCTCAGAGGTGACAGGAGGGGTGAGTACACTATATCCTACCGAGACTTTGTGAA
CATGATGCTGGGGAAACGGTCGGCTGTCTCAAGTTAGTCATGATGTTTGAAGGAAAAGCCA
ACGAGAGCAGCCCCAAGCCAGTTGGCCCCCCTCCAGAGAGAGACATTGCTAGCCTGCCC**TGA**
GGACCCCGCCTGGACTCCCCAGCCTTCCCACCCCATACCTCCCTCCCGATCTTGCTGCCCTT
CTTGACACACTGTGATCTCTCTCTCTCATTGTTTGGTCATTGAGGGTTTGTGTTGTGTTT
TCATCAATGTCTTTGTAAAGCACAAATTATCTGCCTTAAAGGGGCTCTGGGTGCGGGAATCC
TGAGCCTTGGGTCCCCTCCCTCTCTTCTTCCCTCCTTCCCCGCTCCCTGTGCAGAAGGGCTG
ATATCAAACCAAAAACCTAGAGGGGGCAGGGCCAGGGCAGGGAGGCTTCCAGCCTGTGTTCCC
CTCACTTGAGGAACCAGCACTCTCCATCCTTTAGAAAGTCTCCAAGCCAAGTTCAGGCTC
ACTGACCTGGCTCTGACGAGGACCCCAGGCCACTCTGAGAAGACCTTGGAGTAGGGACAAGG
CTGCAGGGCCTCTTTCGGGTTTCTTGGACAGTGCCATGGTTCCAGTGCTCTGGTGTACCCC
AGGACACAGCCACTCGGGGCCCCGCTGCCCCAGCTGATCCCCACTCATTCCACACCTCTTCT
CATCCTCAGTGATGTGAAGGTGGGAAGGAAAGGAGCTTGGCATTGGGAGCCCTTCAAGAAGG
TACCAGAAGGAACCCTCCAGTCCTGCTCTCTGGCCACACCTGTGCAGGCAGCTGAGAGGCAG
CGTGCAGCCCTACTGTCCCTTACTGGGGCAGCAGAGGGCTTCGGAGGCAGAAGTGAGGCCTG
GGGTTTGGGGGGAAAGGTCAGCTCAGTGCTGTTCCACCTTTTAGGGAGGATACTGAGGGGAC
CAGGATGGGAGAATGAGGAGTAAAATGCTCACGGCAAAGTCAGCAGCACTGGTAAGCCAAGA
CTGAGAAATACAAGGTTGCTTGTCTGACCCCAATCTGCTTGAAAAAAAAAAAAAAAAAAAA

MSGELSNRFQGGKAFGLLKARQERRLAEGINREFLCDQKYSDEENLPEKLTAFKEYMEFDLN
NEGEIDLMSLKRMMKELGVPKTHLEMKKMISEVTGGVSDTISYRDFVNMMLGKRSAYLKLVM
MFEGKANESSPKPVGPPPERDIASLP

FIGURE 111A

CGCGCTCCCCGCGCGCCTCCTCGGGCTCCACGCGTCTTGCCCCGCAGAGGCAGCCTCCTCCA
GGAGCGGGGCCCCTGCACACC**ATG**GCCCCCGGGTGGGCAGGGGTGCGCGCCGCCGTGCGCGCC
CGCCTGGCGCTGGCCTTGGCGCTGGCGAGCGTCCTGAGTGGGCCTCCAGCCGTGCGCTGCCC
CACCAAGTGTACCTGCTCCGCTGCCAGCGTGGACTGCCACGGGCTGGGCCTCCGCGCGGTTTC
CTCGGGGCATCCCCGCAACGCTGAGCGCCTTGACCTGGACAGAAATAATATCACCAGGATC
ACCAAGATGGACTTCGCTGGGCTCAAGAACCTCCGAGTCTTGATCTGGAAGACAACCAGGT
CAGCGTCATCGAGAGAGGGCGCCTTCCAGGACCTGAAGCAGCTAGAGCGACTGCGCCTGAACA
AGAATAAGCTGCAAGTCCTTCCAGAATTGCTTTTCCAGAGCACGCCGAAGCTCACCAGACTA
GATTTGAGTGAAAACCAGATCCAGGGGATCCCGAGGAAGGCGTTCCGCGGCATCACCGATGT
GAAGAACCTGCAACTGGACAACAACCACATCAGCTGCATTGAAGATGGAGCCTTCCGAGCGC
TGC GCGATTTGGAGATCCTTACCCTCAACAACAACAACATCAGTCGCATCCTGGTCACCAGC
TTCAACCACATGCCGAAGATCCGAACCTCTGCGCCTCCACTCCAACCACCTCTACTGCGACTG
CCACCTGGCCTGGCTCTCGGATTGGCTGCGACAGCGACGGACAGTTGGCCAGTTCACACTCT
GCATGGCTCCTGTGCATTTGAGGGGCTTCAACGTGGCGGATGTGCAGAAGAAGGAGTACGTG
TGCC CAGCCCCCACTCGGAGCCCCCATCCTGCAATGCCAACTCCATCTCCTGCCCTTCGCC
CTGCACGTGCAGCAATAACATCGTGGACTGTCGAGGAAAGGGCTTGATGGAGATTCTTGCCA
ACTTGCCGGAGGGCATCGTCGAAATACGCCTAGAACAGAACTCCATCAAAGCCATCCCTGCA
GGAGCCTTCACCCAGTACAAGAACTGAAGCGAATAGACATCAGCAAGAATCAGATATCGGA
TATTGCTCCAGATGCCTTCCAGGGCCTGAAATCACTCACATCGCTGGTCCTGTATGGGAACA
AGATCACCGAGATTGCCAAGGGACTGTTTGATGGGCTGGTGTCCCTACAGCTGCTCCTCCTC
AATGCCAACAAGATCAACTGCCTGCGGGTGAACACGTTTCAGGACCTGCAGAACCTCAACTT
GCTCTCCCTGTATGACAACAAGCTGCAGACCATCAGCAAGGGGGCTCTTCGCCCCCTCTGCAGT
CCATCCAGACACTCCACTTAGCCCCAAAACCCATTTGTGTGCGACTGCCACTTGAAGTGGCTG
GCCGACTACCTCCAGGACAACCCCATCGAGACAAGCGGGGCCCGCTGCAGCAGCCCCGCGCCG
ACTCGCCAACAAGCGCATCAGCCAGATCAAGAGCAAGAAGTTCCGCTGCTCAGGCTCCGAGG
ATTACCGCAGCAGGTTCAGCAGCGAGTGCTTCATGGACCTCGTGTGCCCCGAGAAGTGTCGC
TGTGAGGGCACGATTGTGGACTGCTCCAACCAGAAGCTGGTCCGCATCCCAAGCCACCTCCC
TGAATATGTCACCGACCTGCGACTGAATGACAATGAGGTATCTGTTCTGGAGGCCACTGGCA
TCTTCAAGAAGTTGCCAACCTGCGGAAAATAAATCTGAGTAACAATAAGATCAAGGAGGTG
CGAGAGGGAGCTTTCGATGGAGCAGCCAGCGTGCAGGAGCTGATGCTGACAGGGAACCAGCT
GGAGACCGTGCACGGGCGCGTGTTCCGTGGCCTCAGTGGCCTCAAAACCTTGATGCTGAGGA
GTAACCTGATCAGCTGTGTGAGTAATGACACCTTTGCCGGCCTGAGTTCGGTGAGACTGCTG
TCCCTCTATGACAATCGGATCACCACCATCACCCCTGGGGCCTTACCACGCTTGTCTCCCT
GTCCACCATAAACCTCCTGTCCAACCCCTTCAACTGCAACTGCCACCTGGCCTGGCTCGGCA
AGTGGTTGAGGAAGAGGCGGATCGTCAGTGGGAACCCTAGGTGCCAGAAGCCATTTTTCCTC
AAGGAGATTCCCATCCAGGATGTGGCCATCCAGGACTTCACCTGTGATGGCAACGAGGAGAG
TAGCTGCCAGCTGAGCCCGCGCTGCCCCGAGCAGTGACCTGTATGGAGACAGTGGTGCGAT
GCAGCAACAAGGGGCTCCGCGCCCTCCCCAGAGGCATGCCCAAGGATGTGACCGAGCTGTAC
CTGGAAGGAAACCACCTAACAGCCGTGCCCAGAGAGCTGTCCGCCCTCCGACACCTGACGCT
TATTGACCTGAGCAACAACAGCATCAGCATGCTGACCAATTACACCTTCAGTAACATGTCTC
ACCTCTCCACTCTGATCCTGAGCTACAACCGGCTGAGGTGCATCCCCGTCCACGCCTTCAAC
GGGCTGCGGTCCCTGCGAGTGCTAACCCCTCATGGCAATGACATTTCCAGCGTTCCCTGAAGG
CTCCTTCAACGACCTCACATCTCTTTCCCATCTGGCGCTGGGAACCAACCCACTCCACTGTG
ACTGCAGTCTTCGGTGGCTGTGCGAGTGGGTGAAGGCGGGGTACAAGGAGCCTGGCATCGCC
CGCTGCAGTAGCCCTGAGCCCATGGCTGACAGGCTCCTGCTCACCACCCCAACCCACCGCTT
CCAGTGCAAAGGGCCAGTGGACATCAACATTTGTGGCCAAATGCAATGCCTGCCTCTCCAGCC
CGTGCAAGAATAACGGGACATGCACCCAGGACCCTGTGGAGCTGTACCGCTGTGCCTGCCCC

FIGURE 111B

TACAGCTACAAGGGCAAGGACTGCACTGTGCCCATCAACACCTGCATCCAGAACCCCTGTCA
GCATGGAGGCACCTGCCACCTGAGTGACAGCCACAAGGATGGGTTGAGCTGCTCCTGCCCTC
TGGGCTTTGAGGGGCAGCGGTGTGAGATCAACCCAGATGACTGTGAGGACAACGACTGCGAA
AACAAATGCCACCTGCGTGACGGGATCAACAACCTACGTGTGTATCTGTCCGCCTAACTACAC
AGGTGAGCTATGCGACGAGGTGATTGACCACTGTGTGCCTGAGCTGAACCTCTGTCAGCATG
AGGCCAAGTGATCCCCCTGGACAAAGGATTGAGCTGCGAGTGTGTCCCTGGCTACAGCGGG
AAGCTCTGTGAGACAGACAATGATGACTGTGTGGCCCACAAGTGCCGCCACGGGGCCCAGTG
CGTGACACAATCAATGGCTACACATGCACCTGCCCCCAGGGCTTCAGTGGACCCTTCTGTG
AACACCCCCCACCCATGGTCCTACTGCAGACCAGCCCATGCGACCAGTACGAGTGCCAGAAC
GGGGCCCAGTGATCGTGGTGACGAGGAGCCCACCTGCCGCTGCCACCAGGCTTCGCCGG
CCCCAGATGCGAGAAGCTCATCACTGTCAACTTCGTGGGCAAAGACTCCTACGTGGAACCTGG
CCTCCGCCAAGGTCCGACCCAGGCCAACATCTCCCTGCAGGTGGCCACTGACAAGGACAAC
GGCATCCTTCTCTACAAAGGAGACAATGACCCCTGGCACTGGAGCTGTACCAGGGGCCACGT
GCGGCTGGTCTATGACAGCCTGAGTTCCCTCCCAACCACAGTGTACAGTGTGGAGACAGTGA
ATGATGGGCAGTTTCACAGTGTGGAGCTGGTGACGCTAAACCAGACCCTGAACCTAGTAGTG
GACAAAGGAACTCCAAAGAGCCTGGGGAAGCTCCAGAAGCAGCCAGCAGTGGGCATCAACAG
CCCCCTCTACCTTGGAGGCATCCCCACCTCCACCGGCCCTCTCCGCCTTGCGCCAGGGCACGG
ACCGGCCTCTAGGCGGCTTCCACGGATGCATCCATGAGGTGCGCATCAACAACGAGCTGCAG
GACTTCAAGGCCCTCCACACAGTCCCTGGGGGTGTACCAGGCTGCAAGTCTGCACCGT
GTGCAAGCACGGCCTGTGCCGCTCCGTGGAGAAGGACAGCGTGGTGTGCGAGTGCCGCCAG
GCTGGACCGGCCCACTCTGCGACCAGGAGGCCCGGGACCCCTGCCTCGGCCACAGATGCCAC
CATGGAAAATGTGTGGCAACTGGGACCTCATAATGTGCAAGTGTGCCGAGGGCTATGGAGG
GGACTTGTGTGACAACAAGAATGACTCTGCCAATGCCTGCTCAGCCTTCAAGTGTACCATG
GGCAGTGCCACATCTCAGACCAAGGGGAGCCCTACTGCCTGTGCCAGCCCGGCTTTAGCGGC
GAGCACTGCCAACAAGAGAATCCGTGCCTGGGACAAGTAGTCCGAGAGGTGATCCGCCGCCA
GAAAGGTTATGCATCATGTGCCACAGCCTCCAAGGTGCCCATCATGGAATGTGCTGGGGGCT
GTGGGCCCCAGTGCTGCCAGCCCACCCGAGCAAGCGGCGGAAATACGTCTTCCAGTGCACG
GACGGCTCCTCGTTTGTAGAAGAGGTGGAGAGACACTTAGAGTGCGGCTGCCTCGCGTGTTC
C**TAA**GGCCCTGCCCGCCTGCCTGCCACCTCTCGGACTCCAGCTTGATGGAGTTGGGACAGCC
ATGTGGGACCCCTGGTGATTCAGCATGAAGGAAATGAAGCTGGAGAGGAAGGTAAAGAAGA
AGAGAATATTAAGTATATTGTAAATAAACAAAAAATAGAACTAAAAAAAAAAAAAAAAAAAA
AAAAAA

FIGURE 112

MAPGWAGVGAAVRARLALALASVLSGPPAVACPTKCTCSAASVDCHGLGLRAVPRGIPRN
 AERLDLDRNNITRITKMDFAGLKNLRVLHLEDNQSVIERGAFQDLKQLERLRLNKNKLQVL
 PELLFQSTPKLTRLDLSENQIQGIPRKAFRGITDVKNLQLDNNHISCIEDGAFRALRDLEIL
 TLNNNNISRILVTSFNHMPKIRTLRLHSNHLYCDCHLAWLSDWLRQRRTVGQFTLCMAPVHL
 RGFNVADVQKEYVCPAPHSEPPSCNANSISCPSPCTCSNNIVDCRGKGLMEIPANLPEGIV
 EIRLEQNSIKAIPAGAFQYKKLKRIDISKQISDIAPDAFQGLKSLTSLVLYGNKITEIAK
 GLFDGLVSLQLLLLNANKINCLRVNTFQDLQNLNLLSLYDNKLQTISKGLFAPLQSIQTLHL
 AQNPFCVCDCHLKWLDYLDQNPDIETSGARCSSPRRLANKRISQIKSKKFRCSGSEDRSRFS
 SECFMDLVCPEKCRCEGTIVDCSNQKLVRIPSHLPEYVTDLRLNDNEVSVLEATGIFKKLPN
 LRKINLSNNKIKEVREGAFDGAASVQELMLTGNQLETVHGRVFRGLSGLKTLMLRSNLISCV
 SNTDFAGLSSVRLLSLYDNRIITTITPGAFTTLVSLSTINLLSNPFNCNCHLAWLGKWLKRR
 IVSGNPRCQKPFFLKEIPIQDVAIQDFTCDGNEESSCQLSPRCPEQCTCMETVVRC SNKGLR
 ALPRGMPKDVTELYLEGNHLTAVPRELSALRHLLTIDLSNNSISMLTNYTFSNM SHLSTLIL
 SYNRLRCIPVHAFNGLRSLRVLT LHGNDISSVPEGSFNDLTSLSHLALGTNPLHDCSLRWL
 SEWVKAGYKEPGIARCSSPEPMADRLLLLTTPTHRFQCKGPVDINIVAKCNACLSSPCKNNGT
 CTQDPVELYRCACPYSYKGDCTVPINTCIQNPCQHGGTCHLSDSHKDGFS CSCPLGFEGQR
 CEINPDDCEDNDCENNATCVDGINNYVCICPPNYTGELCDEVIDHCVPELNLCQHEAKC IPL
 DKGFSCECVPGYSGLCETDNDDCVAHKCRHGAQCVDTINGYTCTCPQGFS GPFCHEPPPMV
 LLQTSPCDQYECQNGAQCIVVQQEPTCRCPPGFAGPRCEKLITVNFVGKDSYVELASAKVRP
 QANISLQVATDKDNGILLYKGDNDPLALELYQGHVRLVYDSLSSPPTTVYSVETVNDGQFHS
 VELVTNLNQLNLVVDKGTPKSLGKLQKQPAVGINSPLYLGGIPTSTGLSALRQGTDRPLGGF
 HGCIHEVRINNELQDFKALPPQSLGVSPGCKSCTVCKHGLCRSVEKDSVVCECRPGWTGPLC
 DQEARDPCLGHRCHHGKCVATGTSYMCKCAEGYGGDLCDNKND SANACSAFKCHHGQCHISD
 QGEPYCLCQPGFSGEHCQQENPCLGQVVREVIRRQKGYASCATASKVPIMECRGGCGPQCCQ
 PTRSKRRKYVFQCTDGSSFVEEVERHLECGCLACS

Signal peptide:

amino acids 1-27

FIGURE 113

GGATGCAGGACGCTCCCCTGAGCTGCCTGTCACCGACTAGGTGGAGCAGTGTTTCTTCCGCA
GACTCAACTGAGAAGTCAGCCTCTGGGGCAGGCACCAGGAATCTGCCTTTTCAGTTCTGTCT
CCGGCAGGCTTTGAGGATGAAGGCTGCGGGCATTCTGACCCTCATTGGCTGCCTGGTCACAG
GCGCCGAGTCCAAAATCTACACTCGTTGCAAACCTGGCAAAAATATTCTCGAGGGCTGGCCTG
GACAATTACTGGGGCTTCAGCCTTGGAACCTGGATCTGCATGGCATATTATGAGAGCGGCTA
CAACACCACAGCCCCGACGGTCCTGGATGACGGCAGCATCGACTATGGCATCTTCCAGATCA
ACAGCTTCGCGTGGTGCAGACGCGGAAAGCTGAAGGAGAACAACCACTGCCATGTCGCCTGC
TCAGCCTTGATCACTGATGACCTCACAGATGCAATTATCTGTGCCAGGAAAATTGTTAAAGA
GACACAAGGAATGAACTATTGGCAAGGCTGGAAGAAACATTGTGAGGGCAGAGACCTGTCCG
AGTGGA AAAAAGGCTGTGAGGTTTCCTAAACTGGAACCTGGACCCAGGATGCTTTGCAGCAAC
GCCCTAGGATTTGCAGTGAATGTCCAAATGCCTGTGTGCATCTTGTCCCGTTTCCTCCCAATA
TTCCTTCTCAAACCTTGAGAGGGGAAAATTAAGCTATACTTTTAAGAAAATAAATATTTCCAT
TTAAATGTC

FIGURE 114

MKAAGILTLIGCLVTGAESKIYTRCKLAKIFSRAGLDNYWGFSLGNWICMAYYESGYNTTAP
TVLDDGSIDYGIFQINSFAWCRRGKLKENNHCHVACSALITDDLTDALICARKIVKETQGMN
YWQGWKKHCEGRDLSEWKKGCEVS

Signal peptide:

amino acids 1-19

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 115

CAGGCCATTTGCATCCCCTGTCCTTGTGTTTCGGAGCCAGGCCACACCGTCCTCAGCAGTGT
 CATGTGTTAAAAACGCCAAGCTGAATATATC**ATG**CCCCTATTAAAACTTGTACATGGCTCCC
 CATTGGTTTTTTGGAGAAAAGTTCAAGCTTTTTACCTTGGTGTCTGCCTGTATCCCAGTGTTC
 AGGCTGGCTAGACGGCGGAAGAAGATCCTATTTTACTGTCACTTCCCAGATCTGCTTCTCAC
 CAAGAGAGATTCTTTTCTTAAACGACTATACAGGGCCCCAATTGACTGGATAGAGGAATACA
 CCACAGGCATGGCAGACTGCATCTTAGTCAACAGCCAGTTCACAGCTGCTGTTTTTAAGGAA
 ACATTCAAGTCCCTGTCTCACATAGACCCTGATGTCCTCTATCCATCTCTAAATGTCACCAG
 CTTTGACTCAGTTGTTTCTGAAAAGCTGGATGACCTAGTCCCCAAGGGGAAAAAATTCCTGC
 TGCTCTCCATCAACAGATACGAAAGGAAGAAAAATCTGACTTTGGCACTGGAAGCCCTAGTA
 CAGCTGCGTGGAAGATTGACATCCCAAGATTGGGAGAGGGTTCATCTGATCGTGGCAGGTGG
 TTATGACGAGAGAGTCCCTGGAGAATGTGGAACATTATCAGGAATTGAAGAAAATGGTCCAAC
 AGTCCGACCTTGGCCAGTATGTGACCTTCTTGAGGTCTTTCTCAGACAAACAGAAAATCTCC
 CTCCTCCACAGCTGCACGTGTGTGCTTTACACACCAAGCAATGAGCACTTTGGCATTGTCCC
 TCTGGAAGCCATGTACATGCAGTGCCCAGTCATTGCTGTTAATTCGGGTGGACCCTTGGAGT
 CCATTGACCACAGTGTACAGGGTTTCTGTGTGAGCCTGACCCGGTGCACCTTCTCAGAAGCA
 ATAGAAAAGTTCATCCGTGAACCTTCCTTAAAAGCCACCATGGGCCTGGCTGGAAGAGCCAG
 AGTGAAGGAAAAATTTTCCCTGAAGCATTTACAGAACAGCTCTACCGATATGTTACCAAAC
 TGCTGGTA**TAA**TCAGATTGTTTTTAAGATCTCCATTAATGTCATTTTTTATGGATTGTAGACC
 CAGTTTTGAAACCAAAAAAGAAACCTAGAACTAATGCAGAAGAGATCTTTTAAAAAATAAA
 CTTGAGTCTTGAATGTGAGCCACTTTCCTATATACCACACCTCCCTGTCCACTTTTCAGAAA
 AACCATGTCTTTTATGCTATAATCATTCCAAATTTTGCCAGTGTTAAGTTACAAATGTGGTG
 TCATTCCATGTTTCAGCAGAGTATTTTAATTATATTTTCTCGGGATTATTGCTCTTCTGTCTA
 TAAATTTTGAATGATACTGTGCCTTAATTGGTTTTTCATAGTTTAAGTGTGTATCATTATCAA
 AGTTGATTAATTTGGCTTCATAGTATAATGAGAGCAGGGCTATTGTAGTTCCCAGATTCAAT
 CCACCGAAGTGTTCACTGTCATCTGTTAGGGAATTTTGTGTTGTCCTGTCTTTGCCTGGATC
 CATAGCGAGAGTGCTCTGTATTTTTTTTAAGATAATTTGTATTTTTTGCACACTGAGATATAA
 TAAAAGGTGTTTATCATAAAAAATAAAAAAAAAAAAAA

10.1	100%
10.2	100%
10.3	100%
10.4	100%
10.5	100%
10.6	100%
10.7	100%
10.8	100%
10.9	100%
10.10	100%
10.11	100%
10.12	100%
10.13	100%
10.14	100%
10.15	100%
10.16	100%
10.17	100%
10.18	100%
10.19	100%
10.20	100%
10.21	100%
10.22	100%
10.23	100%
10.24	100%
10.25	100%
10.26	100%
10.27	100%
10.28	100%
10.29	100%
10.30	100%
10.31	100%
10.32	100%
10.33	100%
10.34	100%
10.35	100%
10.36	100%
10.37	100%
10.38	100%
10.39	100%
10.40	100%
10.41	100%
10.42	100%
10.43	100%
10.44	100%
10.45	100%
10.46	100%
10.47	100%
10.48	100%
10.49	100%
10.50	100%
10.51	100%
10.52	100%
10.53	100%
10.54	100%
10.55	100%
10.56	100%
10.57	100%
10.58	100%
10.59	100%
10.60	100%
10.61	100%
10.62	100%
10.63	100%
10.64	100%
10.65	100%
10.66	100%
10.67	100%
10.68	100%
10.69	100%
10.70	100%
10.71	100%
10.72	100%
10.73	100%
10.74	100%
10.75	100%
10.76	100%
10.77	100%
10.78	100%
10.79	100%
10.80	100%
10.81	100%
10.82	100%
10.83	100%
10.84	100%
10.85	100%
10.86	100%
10.87	100%
10.88	100%
10.89	100%
10.90	100%
10.91	100%
10.92	100%
10.93	100%
10.94	100%
10.95	100%
10.96	100%
10.97	100%
10.98	100%
10.99	100%
11.00	100%

MPLLLKLVHGSPLVFGGEKFKLFTLVSAIPVFRLARRRKKILFYCHFPDLLLTKRDSFLKRLY
RAPIDWIEEYTTGMADCILVNSQFTAAVFKETFKSLSHIDPDVLYPSLNVTSFDSVVPKELD
DLVPKGKKFLLLSINRYERKKNLTLALEALVQLRGRLTSQDWERVHLIVAGGYDERVLENVE
HYQELKKMVQQSDLGQYVTFLRSSFSDKQKISLLHSCTCVLYTPSNEHFGIVPLEAMYMQCPV
IAVNSGGPLESIDHSVTGFLCEPDVHFSEAIEKFIREPSLKATMGLAGRARVKEKFSPEAF
TEQLYRYVTKLLV

amino acids 1-15

FIGURE 117

GACTACGCCGATCCGAGACGTGGCTCCCTGGGCGGCAGAACC**ATG**TTGGACTTCGCGATCTT
 CGCCGTTACCTTCTTGCTGGCGTTGGTGGGAGCCGTGCTCTACCTCTATCCGGCTTCCAGAC
 AAGCTGCAGGAATTCCAGGGATTACTCCAACCTGAAGAAAAAGATGGTAATCTTCCAGATATT
 GTGAATAGTGGAAGTTTGCATGAGTTCCTGGTTAATTTGCATGAGAGATATGGGCCTGTGGT
 CTCCTTCTGGTTTGGCAGGCGCCTCGTGGTTAGTTTGGGCACTGTTGATGTACTGAAGCAGC
 ATATCAATCCCAATAAGACATCGGACCCTTTTGAAACCATGCTGAAGTCATTATTAAGGTAT
 CAATCTGGTGGTGGCAGTGTGAGTGAAAACCATGAGGAAAAAATTGTATGAAAATGGTGT
 GACTGATTCTCTGAAGAGTAACTTTGCCCTCCTCCTAAAGCTTTCAGAAGAATTATTAGATA
 AATGGCTCTCCTACCCAGAGACCCAGCACGTGCCCCCTCAGCCAGCATATGCTTGGTTTTGCT
 ATGAAGTCTGTTACACAGATGGTAATGGGTAGTACATTTGAAGATGATCAGGAAGTCATTCTG
 CTTCCAGAAGAATCATGGCACAGTTTGGTCTGAGATTGGAAAAGGCTTCTTAGATGGGTCTAC
 TTGATAAAAACATGACTCGGAAAAACAATATGAAGATGCCCTCATGCAACTGGAGTCTGTT
 TTAAGGAACATCATAAAAGAACGAAAAGGAAGGAAGTTCAGTCAACATATTTTTCATTGACTC
 CTTAGTACAAGGGAACCTTAATGACCAACAGATCCTAGAAGACAGTATGATATTTTCTCTGG
 CCAGTTGCATAATAACTGCAAAATTGTGTACCTGGGCAATCTGTTTTTTAACCACCTCTGAA
 GAAGTTCAAAAAAATTATATGAAGAGATAAACCAAGTTTTTTGGAAATGGTCCTGTTACTCC
 AGAGAAAATTGAGCAGCTCAGATATTGTCAGCATGTGCTTTGTGAAACTGTTTCGAACTGCCA
 AACTGACTCCAGTTTCTGCCCAGCTTCAAGATATTGAAGGAAAAATTGACCGATTTATTATT
 CCTAGAGAGACCCTCGTCCTTTATGCCCTTGGTGTGGTACTTCAGGATCCTAATACTTGGCC
 ATCTCCACACAAGTTTGATCCAGATCGGTTTGTATGATGAATTAGTAATGAAAACTTTTTCCT
 CACTTGGATTCTCAGGCACACAGGAGTGTCCAGAGTTGAGGTTTGCATATATGGTGACCACA
 GTACTTCTTAGTGTATTGGTGAAGAGACTGCACCTACTTTCTGTGGAGGGACAGGTTATTGA
 AACAAAGTATGAACTGGTAACATCATCAAGGGAAGAAGCTTGGATCACTGTCTCAAAGAGAT
 AT**TAA**AATTTTATACATTTAAATCATTTGTTAAATTGATTGAGGAAAACAACCATTTAAAAA
 AAATCTATGTTGAATCCTTTTATAAACAGTATCACTTTGTAATATAAACACCTATTTGTAC
 TTAA

FIGURE 118

MLDFAIFAVTFLLALVGAVLYLYPASRQAAGIPGITPTEEKDGNLPDIVNSGSLHEFLVNLH
ERYGPVVSFWFGRRLVSVSLGTVDVLKQHINPNKTSDFETMLKSLLRYQSGGGSVSENHMRK
KLYENGVTDSLKSNFALLLLKLSEELLDKWLSYPETQHVPLSQHMLGFAMKSVTQMVMGSTFE
DDQEVIRFQKNHGTWVSEIGKGFLDGS�DKNMTRKKQYEDALMQLESVLRNIIKERKGRNFS
QHIFIDSLVQGNLNDQQILEDSMIFSLASCIITAKLCTWAICFLTTSSEEVQKKLYEEINQVF
GNGPVTPEKIEQLRYCQHVLCEVTVRTAKLTPVSAQLQDIEGKIDRFIIPRETLVLYALGVVL
QDPNTWPSPHKFDPDRFDDELVMKTFSSLGFSGTQECPELRFAYMVTTVLLSVLVKRLHLLS
VEGQVIETKYELVTSSREEAWITVSKRY

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 271-290

FIGURE 119

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGA
GACCGCCGCCCTTGTCCTCCGAGGGGCC**ATG**GGCCGGGTCTCAGGGCTTGCGCCCTCTCGCTTC
CTGACGCTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACAT
ACAGGCCTGCCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTTCAGCTGG
TGGCCGCGCTCTCTGTACCCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGA
GTCTCCATGTTCAACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGT
GGCCCTGTCCTTCTTCATATTCGAGCGTTGGGAGTGCACTACGTATTGGTACATTTTTGTCT
TCTGCAGTGCCCTTCCAGCTGTCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAA
AAGAAACCCTT**TGA**TTACCTTCATGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGG
CCGCTTCGTATTCCTGGAAGAAGGAAGGCATAGGCTTCGGTTTTCCCCTCGGAAACTGCTTC
TGCTGGAGGATATGTGTTGGAATAATTACGTCTTGAGTCTGGGATTATCCGCATTGTATTTA
GTGCTTTGTAATAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACA
ATTAAAAAAAAAAAAA

FIGURE 120

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLG
LFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVT
EMALFVTVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

Downloaded from www.sciencedirect.com

FIGURE 121

TCCCGGACCCTGCCGCCCTGCCACT**ATG**TCCCGCCGCTCTATGCTGCTTGCCTGGGCTCTCC
CCAGCCTCCTTCGACTCGGAGCGGCTCAGGAGACAGAAGACCCGGCCTGCTGCAGCCCCATA
GTGCCCCGGAACGAGTGGAAGGCCCTGGCATCAGAGTGCGCCCAGCACCTGAGCCTGCCCTT
ACGCTATGTGGTGGTATCGCACACGGCGGGCAGCAGCTGCAACACCCCCGCCTCGTGCCAGC
AGCAGGCCCCGGAATGTGCAGCACTACCACATGAAGACACTGGGCTGGTGCGACGTGGGCTAC
AACTTCCTGATTGGAGAAGACGGGCTCGTATACGAGGGCCGTGGCTGGAAC TTCACGGGTGC
CCACTCAGGTCAC TTATGGAACCCCATGTCCATTGGCATCAGCTTCATGGGCAACTACATGG
ATCGGGTGCCCACACCC CAGGCCATCCGGGCAGCCCAGGGTCTACTGGCCTGCGGTGTGGCT
CAGGGAGCCCTGAGGTCCA ACTATGTGCTCAAAGGACACCGGGATGTGCAGCGTACACTCTC
TCCAGGCAACCAGCTCTACCACCTCATCCAGAATTGGCCACACTACCGCTCCCCC**TG**AGGCC
CTGCTGATCCGCACCCCATTCCTCCCCTCCCATGGCCAAAAACCCCACTGTCTCCTTCTCCA
ATAAAGATGTAGCTC

FIGURE 122

MSRRSMLLAWALPSLLRLGAAQETEDPACCSPIVPRNEWKALASECAQHLSLPLRYVVVSHT
AGSSCNTPASCCQQQARNVQHYHMKTLGWCDVGYNFLIGEDGLVYEGRGWNFTGAHSGHLWNP
MSIGISFMGNVMDRVPTPQAIRAAQGLLACGVAQGALRSNYVLKGHRDVQRTLSPGNQLYHL
IQNWPHYRSP

Signal peptide:

amino acids 1-20

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted May 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 123

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAG**ATG**CAACT
 GACTCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGG
 ATGATGGTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCC
 CGGGTGCCTCGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCT
 CCTAGGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCCTTGGGCAGCCCCCAACCGCC
 CGAACCACAGCCCCCACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTAC
 TCCAACATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAA
 TGGGACCTTCAGCGTCCACTTCCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCC
 TCGTGCCCCCAGTAAAGCTGTAGAGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAG
 GCCTCCAAAATCTTCAACTGCCGGATGGAGTGGGAGAAGGTAGAACGGGGCCGCCGGACCTC
 GCTTTGCACCCACGACCCAGCCAAGATCTGCTCCCGAGACCACGCTCAGAGCTCAGCCACCT
 GGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTAT
 CGGCTGGTCCAGAAGGTGTGCCAGATTACAACTACCATAGTGATACCCCTACTACCCATC
 TGGG**TGA**CCCCGGGGCAGGCCACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCCATGC
 AGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTCAGGCAGGGAGGGGGGTGGAG
 ACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAGAGAAAGGGTCCCAAGTG
 CTGGTCCCAACCTGAAGCTGTGGAGTGAAGTAGATCACAGGAGCACTGGAGGAGGAGTGGGCT
 CTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGGTCCCCGAGGCC
 TGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGTTT
 CTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACAGCCTG
 TCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGGGC
 CAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCTTGTCGTGTGCTGAGCATG
 GCATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTC
 CAGCCAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATT
 GCTGATGGCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCACACTAAGGCC
 ACAGCCCATCCGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAG
 CATCCATGTCCCGGAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTCTCC
 CGGATCTGGATGGCGCCGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGA
 GCATGTGCTGGATCTGTTCTGTGTGTCTGTCTGTGGGTGGGGGGAGGGGAGGGAAGTCTTGT
 GAAACCGCTGATTGCTGACTTTTGTGTGAAGAATCGTGTTCTTGAGCAGGAAATAAGCTT
 GCCCCGGGGCA

FIGURE 124

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66521
><subunit 1 of 1, 252 aa, 1 stop
><MW: 28127, pI: 8.91, NX(S/T): 5
MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMAN
STLLGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVD
HGNGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGR
RTSLCTHDPAKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPY
YPSG
```

Important features of the protein:**Signal peptide:**

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 125

GTGAATGTGAGGGTTTGATGACATTTTCAGATGTCTAGGAACCCAGAGTGGGTGCAGGGGCCCA
GGCAGGGCTGATTCTTGGGCGGAGGAGAGTAGGGTAAAGGGTTCTGCATGAGCTCCTTAAAG
GACAAAGGTAACAGAGCCAGCGAGAGAGCTCGAGGGGAGACTTTGACTTCAAGCCACAGAAT
TGGTGGAAAGTGTGCGCGCCGCCGCCGCTCGCTCCTGCAGCGCTGTGACCTAGCCGCTAG
CATCTTCCCGAGCACCGGGATCCCGGGGTAGGAGGCGACGCGGGCGAGCACCAGCGCCAGCC
GGCTGCGGCTGCCCACACGGCTCACC**ATG**GGCTCCGGGCGCCGGGCGCTGTCCGCGGTGCC
GCCGTGCTGCTGGTCTCACGCTGCCGGGGCTGCCCGTCTGGGCACAGAACGACACGGAGCC
CATCGTGCTGGAGGGCAAGTGTCTGGTGGTGTGCGACTCGAACCCGGCCACGGACTCCAAGG
GCTCCTCTTCTCCCCGCTGGGGATATCGGTCCGGGCGGCCAACTCCAAGGTCGCCTTCTCG
GCGGTGCGGAGCACCAACCACGAGCCATCCGAGATGAGCAACAAGACGCGCATCATTTACTT
CGATCAGATCCTGGTGAATGTGGGTAATTTTTTTCACATTGGAGTCTGTCTTTGTAGCACCAA
GAAAAGGAATTTACAGTTTCAGTTTTTCACGTGATTAAAGTCTACCAGAGCCAACTATCCAG
GTTAACTTGATGTTAAATGGAAAACCAGTAATATCTGCCTTTGCGGGGGACAAAGATGTTAC
TCGTGAAGCTGCCACGAATGGTGTCTGCTCTACCTAGATAAAGAGGATAAGGTTTACCTAA
AACTGGAGAAAGGTAATTTGGTTGGAGGCTGGCAGTATTCACGTTTTCTGGCTTTCTGGTG
TTCCCCCTA**TAG**GATTCAATTTCTCCATGATGTTTCATCCAGGTGAGGGATGACCCACTCCTG
AGTTATTGGAAGATCATTTTTTTCATCATTGGATTGATGTCTTTTATTGGTTTCTCATGGGTG
GATATGGATTCTAAGGATTCTAGCCTGTCTGAACCAATACAAAATTTACAGATTATTTGTG
TGTGTCTGTTTCAGTATATTTGGATTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTA
ACAGTCAAAAGCTGTCTGCAAGACTTATTCTGAATTTCAATTCCTGGGATTACTGAATTAGT
TACAGATGTGGAATTTTATTTGTTTAGTTTTTAAAAGACTGGCAACCAGGTCTAAGGATTAGA
AACTCTAAAGTTCTGACTTCAATCAACGGTTAGTGTGATACTGCCAAAGAACTGTATACTG
TGTTAATATATTGATTATATTTGTTTTTATTCCTTTGGAATTAGTTTGTTTGGTTCTTGTA
AAAACCTGGATTTTTTTTTTTCAGTAACTGGTATTATGTTTTTCTCTTAAATAAGGTAATGAA
TGGCTTGCCACAAATTTACCTTGACTACGATATCATCGACATGACTTCTCTCAAAAAAAAA
GAATGCTTCATAGTTGTATTTTAATTGTATATGTGAAAGAGTCATATTTTCCAAGTTATATT
TTCTAAGAAGAAGATAGATCATAAATCTGACAAGGAAAAAGTTGCTTACCCAAAATCTAAG
TGCTCAATCCCTGAGCCTCAGCAAAACAGCTCCCCCTCCGAGGGAAATCTTATACTTTATTGC
TCAACTTTAATTAAAATGATTGATAATAACCACTTTATTAAAAACCTAAGGTTTTTTTTTTT
TCCGTAGACATGACCACTTTATTAACCTGGTGGTGGGATGCTGTTGTTTCTAATTATACCTAT
TTTTCAAGGCTTCTGTTGTATTTGAAGTATCATCTGGTTTTGCCTTAACTCTTTAAATTGTA
TATATTTATCTGTTTAGCTAATATTAATTCAAATATCCCATATCTAAATTTAGTGCAATAT
CTTGTCTTTTGTATAGGTCATATGAATTCATAAAATTATTTATGTCTGTTATAGAATAAAGA
TTAATATATGTTAAAAAAA

127/249

FIGURE 126

MSGRRALSAVPAVLLVLTLPGLPVWAQNDTEPIVLEGKCLVVCDSNPATDSKGSSSSPLGI
SVRAANSKVAFSAVRSTNHEPSEMSNKTRIIYFDQILVNVGNFFTLESVVFVAPRKGIYSFSF
HVIKVYQSQTIQVNLMLNGKPVISAFAGDKDVTREAAATNGVLLYLDKEDKVYLKLEKGNLVG
GWQYSTFSGFLVFPL

Signal peptide:

amino acids 1-27

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 127

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTC
GCCCTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGC
TTTCTTCTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTA
TTGACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTC
TATATCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTT
GAAGAGTATAAACCCAGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCT
TGGGCTTTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGG
CCAGGCACAGTGGGCATTTCATGGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGAC
GCTGGTCATTATCTTGCTGCATGTATTCTGGGGCATTGTATTTTTTTGATGGCTGTGAGAAGA
AAAAGTGGGGCATCCTCCTTATCGTTCTCCTGACCCACCTGCTGGTGTGAGCCAGACCTTC
ATAAGTTCTTATTATGGAATAAACCTGGCGTCAGCATTTATAATCCTGGTGCTCATGGGCAC
CTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTCTGCCAAG
ACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGATAACCTCAGGGAACCAGCACTTCCCAA
ACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCTTTTTCTGAAAATCCCTTTTTCTG
GTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 128

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66658
><subunit 1 of 1, 257 aa, 1 stop
><MW: 28472, pI: 9.33, NX(S/T): 0
MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDN
KDGPTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLG
GIMSGVFSFVNTLSDSLGPGTVGIHGDSPOFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKKW
GILLIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGTWAFILAAGGSCRSLKCLLCQDKN
FLLYNQRSR
```

Important features of the protein:**Signal peptide:**

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 129

CGGCAACCAGCCGCCGCCACCACCGCTGCCACTGCCGCCCTGCCGGGGCC**ATG**TTTCGCTCTGGGCTTGCCCTTCT
 TGGTGCTCTTGGTGGCCTCGGTTCGAGAGCCATCTGGGGGTTCTGGGGCCCAAGAACGTCTCGCAGAAAAGACGCCG
 AGTTTTCGAGCGCACCTACGTGGACGAGGTCAACAGCGAGCTGGTCAACATCTACACCTTCAACCATACTGTGACCC
 GCAACAGGACAGAGGGCGTGGTGTGTCTGTGAACGTCTGAACAAGCAGAAGGGGGCGCCGTTGCTGTTTGTGG
 TCCGCCAGAAGGAGGCTGTGGTGTCTTCCAGGTGCCCTAATCCTGCGAGGGATGTTTCAGCGCAAGTACCTCT
 ACCAAAAAGTGAACGAACCTGTGTGAGCCCCCACCAGAATGAGTCGGAGATTAGTTCTTCTACGTGGATG
 TGTCCACCCTGTACCAAGTCAACACCACATACCAGTCCGGGTGAGCCGATGGACGATTTTGTGCTCAGGACTG
 GGGAGCAGTTTCAGCTTCAATACCACAGCAGCACAGCCAGTACTTCAAGTATGAGTTCCCTGAAGGCGTGGACT
 CGGTAATTGTCAAGGTGACCTCCAACAAGGCCTTCCCTGCTCAGTCATCTCCATTTCAGGATGTGCTGTGTCTG
 TCTATGACCTGGACAACAACGTAGCCTTTCATCGGCATGTACCAGACGATGACCAAGAAGGCGGCCATCACCGTAC
 AGCGCAAAGACTTCCCCAGCAACAGCTTTTATGTGGTGGTGGTGGTGAAGACCGAAGACCAAGCCTGCGGGGGCT
 CCCTGCCTTTTACCCCTTTCGAGAAAGATGAACCGGTGATCAAGGGCACCGCCAGAAAACCCTGTGAGTGTGG
 TGTCTCAAGCAGTCAGTCTGAGGCATACGTGAGTGGGATGCTCTTTTGCCTGGGTATATTTCTCTCTCTTTTACC
 TGCTGACCGTCCCTCGCCTGGCCTGCTGGGAGAAGTGGAGGCAGAAGAAGAAGACCCTGCTGGTGGCCATTGACCGAG
 CCTGCCAGAAAAGCGGTACCCCTCGAGTCTGGCTGATTCTTTTCCCTGGCAGTTCCCTTATGAGGGTTACAACT
 ATGGCTCCTTTGAGAATGTTTCTGGATCTACCGATGGTCTGGTTGACAGCGCTGGCACTGGGGACCTCTCTTACG
 GTTACAGGGGCGCTCCTTTGAACCTGTAGGTACTCGGCCCGAGTGGACTCCATGAGCTCTGTGGAGGAGGATG
 ACTACGACACATTGACCGACATCGATTCCGACAAGAATGTGATTCGCACCAAGCAATACCTCTATGTGGCTGACC
 TGGCAGGGAAGGACAAGCGTGTCTGCGGAAAAAGTACCAGATCTACTTCTGGAACATTGGCCACCATTGCTGTCT
 TCTATGCCCTTCTGTGGTGCAGTGGTGTGATCACCTACCAGACGGTGGTGAATGTGACAGGGAATCAGGACATCT
 GCTACTACAACCTTCTGTGCGCCACCCACTGGGCAATCTCAGCGCCTTCAACAACATCCTCAGCAACCTGGGGT
 ACATCCTGCTGGGGCTGCTTTTCTGCTCATCATCCTGCAACGGGAGATCAACCACAACCGGGCCCTGCTGCGCA
 ATGACCTCTGTGCCCTGGAATGTGGGATCCCCAACACTTTGGGCTTTTCTACGCCATGGGCACAGCCCTGATGA
 TGGAGGGGCTGCTCAGTGTGCTATCATGTGTGCCCAACTATACCAATTTCCAGTTTGACACATCGTTTATGT
 ACATGATCGCCGGACTCTGCATGCTGAAGCTCTACCAGAAGCGGCACCCGGACATCAACGCCAGCGCTACAGTG
 CCTACGCCCTGCCTGGCCATTGTGATCTTCTCTGTGCTGGGCGTGGTCTTTGGCAAAGGGAACACGGCGTTCT
 GGATCGTCTTCTCCATCATTCACATCATCGCCACCCTGCTCCTCAGCACGCAGCTCTATTACATGGGCCGGTGA
 AACTGGACTCGGGGATCTTCCGCCGATCCTCCACGTGCTCTACACAGACTGCATCCGGCAGTGCAGCGGGCCGC
 TCTACGTGGACCGCATGGTGTGCTGGTGCATGGGCAACGTCACTGGTGGTGGCTGCCTATGGGCTTATCA
 TGCGCCCCAATGATTTTCGCTTCTCTACTTGTGGCCATTGGCATCTGCAACCTGCTCCTTTACTTTCGCTTCTACA
 TCATCATGAAGCTCCGGAGTGGGAGAGGATCAAGCTCATCCCCCTGCTCTGCATCGTTTGCACCTCCGTGGTCT
 GGGGCTTCGCGCTCTTCTTCTTCTCCAGGACTCAGCACCTGGCAGAAAACCCCTGCAGAGTCGAGGGAGCACA
 ACCGGGACTGCATCCTCCTCGACTTCTTTGACGACCAGCAGATCTGGCACTTCTCTCCTCCATCGCCATGTTCC
 GGTCTTCTCTGGTGTGCTGACACTGGATGACGACCTGGATACTGTGACGCGGGACAAGATCTATGTCTTCT**TAGC**
 AGGAGCTGGGCCCTTCGCTTACCTCAAGGGGCCCTGAGCTCCTTTGTGTATAGACCGGTCACTCTGTCTGTCT
 GTGGGGATGAGTCCCAGCACCGCTGCCAGCACTGGATGGCAGCAGGACAGCCAGGTCTAGCTTAGGCTTGGCCT
 GGGACAGCCATGGGGTGGCATGGAACCTTGCAGCTGCCCTCTGCCGAGGAGCAGGCCTGCTCCCCTGGAACCCCC
 AGATGTTGGCCAAATTGCTGCTTTCTTCTCAGTGTGGGGCTTCCATGGGCCCTGTCTTTGGCTCTCCATTT
 GTCCCTTTGCAAGAGGAAGGATGGAAGGGACACCTCCCCATTTTCATGCCCTTGCATTTTGGCCGCTCCTCCCC
 ACAATGCCCCAGCCTGGGACCTAAGGCCTCTTTTCTCCCTACTCCCACTCCAGGGCCTAGTCTGGGGCCTGA
 ATCTCTGTCTGTATCAGGGCCCCAGTCTCTTTGGGCTGTCCCTGGCTGCCATCACTGCCATTCCAGTCAGCC
 AGGATGGATGGGGGTATGAGATTTTGGGGGTGGCCAGCTGGTGGCAGACTTTTGGTGTAAAGCCTGCAAGGGG
 CCTGGGGCAGTGGTATTTCTTCCCTCTGACCTGTGCTCAGGGCTGGCTCTTTAGCAATGCGCTCAGCCCAATT
 TGAGAACCGCCCTTCTGATTCAAGAGGCTGAATTCAGAGGTCACCTCTTCATCCCATCAGCTCCCAGACTGATGCC
 AGCACCAGGACTGGAGGGAGAAGCGCCTCACCCCTTCCCTTCTTCTTCCAGGCCCTTAGTCTTGCCAAACCCC
 AGCTGGTGGCCTTTCAGTGCCATTGACACTGCCCAAGAATGTCCAGGGGCAAAGGAGGGATGATACAGAGTTCAG
 CCCGTTCTGCCTCCACAGCTGTGGGCACCCAGTGCTTACCTTAGAAAGGGGCTTCAGGAAGGGATGTGCTGTTT
 CCTCTACGTGCCAGTCCTAGCCTCGCTCTAGGACCCAGGGCTGGCTTCTAAGTTTCCGTCCAGTCTTCAGGCA
 AGTTCTGTGTAGTCATGCACACATACCTATGAAACCTTGGAGTTTACAAAGAATTGCCCCAGCTCTGGGCAC
 CCTGGCCACCCTGGTCTTGGATCCCCCTCGTCCCACCTGGTCCACCCAGATGCTGAGGATGGGGGAGCTCAGG
 CGGGGCTCTGCTTTGGGGATGGGAATGTGTTTTTCTCCAAACTTGTTTTTATAGCTCTGCTTGAAGGGCTGGG
 AGATGAGGTGGTCTGGATCTTTTCTCAGAGCGTCTCCATGCTATGGTTGCATTTCCGTTTTCTATGAATGAATT
 TGATTTCAATAAACCAACCAGACTCAAAAAAAAAAAAAA

FIGURE 130

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66659

><subunit 1 of 1, 832 aa, 1 stop

><MW: 94454, pI: 6.94, NX(S/T): 12

MFALGLPFLVLLVASVESHGLGVLGPKNVSQKDAEFERTYVDEVNSELVNIYTFNHTVTRNRT
EGVRVSVNVLNKQKGAPLLFVVRQKEAVVSFQVPLILRGMFQRKYLYQKVERTLCQPPTKNE
SEIQFFFYVDVSTLSPVNTTYQLRVSRMDDFVLRGTGEQFSFNTTAAQPQYFKYEFPEGVDSVI
VKVTSNKAFFPCSVISIQDVLCPVYDLNNAVAFIGMYQTMTKKAAITVQRKDFPSNSFYVVVV
VKTEDQACGGSPLPFYPFAEDEPVDQGHRQKTL SVLV SQAVTSEAYVSGMLFCLGIFLSFYLL
TVLLACWENWRQKKKTLLVAIDRACPESGHPRVLADSFPGSSPYEGYNYGSFENVSGSTDGL
VDSAGTGDLSYGYQGRSFEPVGTRPRVDSMSSVEEDDYDTLTDIDSDKNVIRTKQYLYVADL
ARKDKRVLRRKKYQIYFWNIATIAVFYALPVVQLVITYQTVVNVVTGNQDICYYNFLCAHPLGN
LSAFNNILSNLGYILLGLLFLLIILQREINHNRRALLRNDLCALECGIPKHFGFLFYAMGTALM
MEGLLSACYHVCPNYTNFQFDTSFMYMIAGLCMLKLYQKRHPDINASAYSAYACLAIVIFFS
VLGVVFGKGNTAFWIVFSIIHIIATLLLSTQLYYMGRWKLDSGIFRRILHVLYTDCIRQCSG
PLYVDRMVLLVMGNVINWSLAAYGLIMRPNDFASYLLAIGICNLLLYFAFYIIMKLRSGERI
KLIPLLCIVCTSVVWGFALFFFFQGLSTWQKTPAESREHNRDCILLDDFFDDHDIWHFLSSIA
MFGSFLVLLTLDDDLDTVQRDKIYVF

Important features of the protein:

Signal peptide:

amino acids 1-18

Transmembrane domains:

amino acids 292-317, 451-470, 501-520, 607-627, 751-770

Leucine zipper pattern.

amino acids 497-518

N-glycosylation sites.

amino acids 27-30, 54-57, 60-63, 123-126, 141-144, 165-168, 364-367, 476-479, 496-499, 572-575, 603-606, 699-702

FIGURE 131

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGC
 TCTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTG
 ACC**ATG**GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGC
 CCAGCCTGCAGAGCTGTCTGTGGAAGTTCCAGAAAACATATGGTGGAATTTCCCTTTATACC
 TGACCAAGTTGCCGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTGAGGGGACTCA
 GGCAAGGCAACTGAGGGGCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAG
 GGCCCTGGACCGAGAGGAGCAGGCAGAGTACCAGCTACAGGTACCCCTGGAGATGCAGGATG
 GACATGTCTTGTGGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTG
 CCCCATTTCTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCC
 CTTCTCTTCTTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACTCGGATCTTCGAT
 TCCACATCCTGAGCCAGGCTCCAGCCCAGCCTTCCCCAGACATGTTCCAGCTGGAGCCTCGG
 CTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGCACCAGCCTTGACCACGCCCTGGAGAGGAC
 CTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACCAGGCCACTG
 CCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGTCCCTAGAGCCTATCCACCTGGCA
 GAGAATCTCAAAGTCTATACCCGCACCACATGGCCAGGTACACTGGAGTGGGGGTGATGT
 GCACTATCACCTGGAGAGCCATCCCCCGGGACCCCTTTGAAGTGAATGCAGAGGGAAACCTCT
 ACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCCAGGTGCGGGCTCAG
 AATTCCTATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTGATGGATGAGAA
 TGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCTCAGTCCAC
 CAGGTACTGAAGTGAAGTACTGTGAGCAGAGGATGCAGATGCCCCCGGCTCCCCCAATTCC
 CACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTCCA
 GGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAAC
 TCCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGT
 GAAGTGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCCAGAT
 TGGGCCTATAAGCCTCCCTGAGGATGTGGAGCCCGGGACTCTGGTGCCATGCTAACAGCCA
 TTGATGCTGACCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACA
 GAAGGGACTTTTGGCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAA
 GAACCTCAGTTATGAGGCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGA
 AGCTGGTGGGGCCAGGCCCAGGCCCTGGAGCCACCGCCACGGTGAAGTGTGCTAGTGGAGAGA
 GTGATGCCACCCCCAAGTTGGACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCC
 AGCCGGCTCTTTCTGCTGACCATCCAGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTCTT
 CCTAGTCAATGACTCAGAGGGCTGGCTCTGCATTGAGAAATTTCTCCGGGGAGGTGCACACC
 GCCCAGTCCCTGCAGGGCGCCAGCCTGGGGACACCTACACGGTGCTTGTGGAGGGCCAGGA
 TACAGCCCTGACTCTTGCCCTGTGCCCTCCCAATACCTCTGCACACCCCGCCAAGACCATG
 GCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCTGGCCAGTGGGCACGGTCCCTACAGC
 TTCACCCCTTGGTCCCAACCCACGGTGCAACGGGATTGGCGCCTCCAGACTCTCAATGGTTC
 CCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGTGAACACATAATCCCCGTGG
 TGGTCAGCCACAATGCCCAGATGTGGCAGCTCCTGGTTCGAGTGATCGTGTGCTGCAAC
 GTGGAGGGGCAGTGCATGCGCAAGGTGGGCCGCATGAAGGGCATGCCACGAAGCTGTGCGC
 AGTGGGCATCCTTGTAGGCACCCCTGGTAGCAATAGGAATCTTCCTCATCTCATTTTACCC
 ACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCCCTGAAGGCG
 ACTGTCT**TGA**ATGGCCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAGTCCC
 CTGGGAGAGAGCCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCCCTCCAT
 CTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCAAC
 TTTATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTTTGCCCAATAATAAAGCCCCA
 GAGAACTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 132

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGD SG
KATEGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVP
HFSQAIYRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSDFQLEPRL
GALALSPKGSTSLDHALERTYQLLVQVKMDQASGHQATATVEVSIIESTWVSLEPIHLAE
NLKVLYPHHMAQVHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQN
SHGEDYAAPLELHVLVMDENDNPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSH
VVYQLLSPEPEDGVEGRAFQVDPTSGSVTLGVLP L RAGQNILLVLAMDLAGAEGGFSSTCE
VEVAVTDINDHAPEFITSQIGPISLPEDVEPGTLVAMLT AIDADLEPAFRLMDFAIERGDTE
GTFGLDWEFDSGHVRLRLCKNLSYEAAPSHEVVVVVQSVAKLVGPGPGPGATATVTVLVERV
MPPPKLDQESYEASVPISAPAGSFLLTIQPSDPISR TLRFSLVNDSEGWLCIEKFSGEVHTA
QSLQGAQPGDTYTVLVEAQDTALT LAPVPSQYLCTPRQDHGLIVSGPSKDPDLASGHGPYSF
TLGPNPTVQRDWRLQTLNGSHAYLT LALHWWVEPREHIIPVVVSHNAQMWQLLVRVIVCRCNV
EGQCMRKVGRMKGMPTKLSAVGILVGT LVAIGIFLILIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 133

CCGGGGAC**ATG**AGGTGGGATATACTGTTCAATTGGGGCCCTTATTGGGTCCAGCATCTGTGGGCCAA
GAAAAATTTTTTGGGGACCAAGTTTTGAGGATTAATGTCAGAAATGGAGACGAGATCAGCAA
ATTGAGTCAACTAGTGAATTCAAACAACCTTGAAGCTCAATTTCTGGAAATCTCCCTCCTCCT
TCAATCGGCCCTGTGGATGTCCTGGTCCCATCTGTCAGTCTGCAGGCATTTAAATCCTTCCTG
AGATCCCAGGGCTTAGAGTACGCAGTGACAATTGAGGACCTGCAGGCCCTTTTAGACAATGA
AGATGATGAAATGCAACACAATGAAGGGCAAGAACGGAGCAGTAATAACTTCAACTACGGGG
CTTACCATTCCCTGGAAGCTATTTACCACGAGATGGACAACATTGCCGCAGACTTTTCCTGAC
CTGGCGAGGAGGGTGAAGATTGGACATTCGTTTGAAAACCGGCCGATGTATGTACTGAAGTT
CAGCACTGGGAAAGGCGTGAGGCGGCCGCGCGTTTGGCTGAATGCAGGCATCCATTCCCGAG
AGTGGATCTCCCAGGCCACTGCAATCTGGACGGCAAGGAAGATTGTATCTGATTACCAGAGG
GATCCAGCTATCACCTCCATCTTGAGAGAAAATGGATATTTTCTTGTTGCCTGTGGCCAATCC
TGATGGATATGTGTATACTCAAACCTCAAACCGATTATGGAGGAAGACGCGGTCCCGAAATC
CTGGAAGCTCCTGCATTGGTGCTGACCCAAATAGAACTGGAACGCTAGTTTTGCAGGAAAG
GGAGCCAGCGACAACCCTTGCTCCGAAGTGTACCATGGACCCACGCCAATTCGGAAGTGGA
GGTGAAATCAGTGGTAGATTTTCATCCAAAACATGGGAATTTCAAGGGCTTCATCGACCTGC
ACAGCTACTCGCAGCTGCTGATGTATCCATATGGGTACTCAGTCAAAAAGGCCCCAGATGCC
GAGGAACTCGACAAGGTGGCGAGGCTTGCGGCCAAAGCTCTGGCTTCTGTGTCGGGCACTGA
GTACCAAGTGGGTCCCACCTGCACCACCTGTCTATCCAGCTAGCGGGAGCAGCATCGACTGGG
CGTATGACAACGGCATCAAATTTGCATTACATTTGAGTTGAGAGATACGGGGACCTATGGC
TTCCTCCTGCCAGCTAACCAGATCATCCCCACTGCAGAGGAGACGTGGCTGGGGCTGAAGAC
CATCATGGAGCATGTGCGGGACAACCTCTAC**TAG**GCGATGGCTCTGCTCTGTCTACATTTAT
TTGTACCCACACGTGCACGCACTGAGGCCATTGTTAAAGGAGCTCTTTCCTACCTGTGTGAG
TCAGAGCCCTCTGGGTTTGTGGAGCACACAGGCCTGCCCCCTCTCCAGCCAGCTCCCTGGAGT
CGTGTGTCCTGGCGGTGTCCCTGCAAGAACTGGTTCCTGCCAGCCTGCTCAATTTTGGTCCTG
CTGTTTTTGTATGAGCCTTTTGTCTGTTTCTCCTTCCACCCTGCTGGCTGGGCGGCTGCACTC
AGCATACCCCTTCTGGGTGGCATGTCTCTCTCTACCTCATTTTTTAGAACCAAAGAACATC
TGAGATGATTCTCTACCCTCATCCACATCTAGCCAAGCCAGTGACCTTGCTCTGGTGGCACT
GTGGGAGACACCACCTTGTCTTTAGGTGGGTCTCAAAGATGATGTAGAATTTCTTTAATTTT
TCGCAGTCTTCCCTGGAAAATATTTTTCTTTTGAGCAGCAAATCTTGATAGGATATCAGTGAAG
GTCTCTCCCTCCCTCCTCTCCTGTTTTTTTTTTTTTTTGGAGACAGAGTTTTGCTCTTGTTGCC
CAGGCTGGAGTGTGATGGCTCGATCTTGGCTCACCACAACCTCTGCCTCCTGGGTTCAAGCA
ATTCTCCTGCCTCAGCCTCTTGAGTAGCTTGGTTTATAGGCGCATGCCACCATGCCTGGCTA
ATTTTGTGTTTTTAGTAGAGACAGGGTTTCTCCATGTTGGTCAGGCTGGTCTCAAACCTCCCA
ACCTCAGGTGATCTGCCCTCCTTGGCCTCCCAGAGTGCTGGGATTACAGGTGTGAGCCACTG
TGCCGGGGCCCGTCCCCTCCTTTTTTAGGCCTGAATACAAAGTAGAAGATCACTTTCCTTCAC
TGTGCTGAGAATTTCTAGATACTACAGTTCTTACTCCTCTCTTCCCTTTGTTATTTCAGTGTG
ACCAGGATGGCGGGAGGGGATCTGTGTCACTGTAGGTACTGTGCCCAGGAAGGCTGGGTGAA
GTGACCATCTAAATTGCAGGATGGTGAAATTATCCCCATCTGTCCTAATGGGCTTACCTCCT
CTTTGCCTTTTGAACTCACTTCAAAGATCTAGGCCTCATCTTACAGGTCTTAAATCACTCAT
CTGGCCTGGATAATCTCACTGCCCTGGCACATTCCTATTTGTGCTGTGGTGTATCCTGTGTT
TCCTTGTCTGGTTTTGT
TCTGTCTATTTTGTATCCTGGACCACAAGTTCCTAAGTAGAGCAAGAATTCATCAACCAGCT
GCCTCTTGTTTCATTTACCTCAGCACGTACCATCTGTCCTTTTGTGTTGTTGTTGTTTGT
TTGTTTTTTTTGCTTTTACCAAACATGTCTGTAAATCTTAACCTCCTGCCTAGGATTTGTACA
GCATCTGGTGTGTGCTTATAAGCCAATAAATATTCAATGTGAAAAAAAAAAAAAAAAAAAA

FIGURE 134

MRWILFIGALIGSSICGQEKFFGDQVLRINVRNGDEISKLSQLVNSNNLKLNFWKSPSSFNR
PVDVLVPSVSLQAFKSFLRSQGLEAYVTIEDLQALLDNEDDEMQHNEGQERSNNFNYGAYH
SLEAIYHEMDNIAADFPDLARRVKIGHSFENRPMYVLKFSTGKGVRRPAVWLNAGIHSREWI
SQATAIWTARKIVSDYQRDPAITSILEKMDIFLLPVANPDGYVYTQTQNRLWRKTRSRNPGS
SCIGADPNRNWNASFAGKGASDNPCSEVYHGPHANSEVEVKSVVDFIQKHGNFKGFIDLHSY
SQLLMYPYGYSVKKAPDAEELDKVARLAAKALASVSGTEYQVGPTCTTVYPASGSSIDWAYD
NGIKFAFTFELRDTGTYGFLLPANQIIPTAEETWLGLKTIMEHVRDONLY

Signal peptide:

amino acids 1-16

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

FIGURE 136

MASYLYGVLEFAVGLCAPIYCVSPANAPSAYPRPSSTKSTPASQVYSLNTDFAFRLYRRLVLE
TPSQNIFFSPVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHSLTVP
SKDLTLKMGSALFVKKELQLQANFLGNVKRLYEAEVFSTDFSNPSIAQARINSHVKKKTQ GK
VVDIIQGLDLLTAMVLVNHIFFKAKWEKPFHLEYTRKNFPFLVGEQVTVQVPMMHQKEQFAF
GVDTELNCFVLQMDYKGDVAFFVLPSKGKMRQLEQALSARTLIKWSHSLQKRWIEVFIPRF
SISASYNLETILPKMGIQNAFDKNADFSGIAKRDSLQVSKATHKAVLDVSEEGTEATAATTT
KFIVRSKDGPSYFTVSFNRTFLMMITNKATDGILFLGKVENPTKS

Signal peptide:

amino acids 1-20

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 137

[illegible]

FIGURE 138

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATI
SGSSVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSE
SSTPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSS
TTSSGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTT
SNGAGTATNSESSTTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSS
GASTATNSDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGA
NTATNSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEAST
ATNSESSTVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAV
SEAKPGGSLVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGP
GPGGNHGAPHRPRWSPNWFWRPVS SIAMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 139

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTC
CCAGCAAT**ATG**CATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCC
CTGTCTGGATGGGCGGCCAGCGATGACCCCATTGAGAAGGTCATTGAAGGGATCAACCGAGG
GCTGAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGGAATCACGC
ATGCCGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGC
AAGGAGTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGAT
CAACCATGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACG
CTGCTGGACAGGCCGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCAC
CAGGCTGGGAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGG
AAAGGAAGTGGAGAAGCTTGGCCAAGGTGCCCAACCATGCTGCTGGCCAGGCCGGAAGGAGC
TGCAGAATGCTCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGC
AACCATCAAAGCGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGG
GGCCTCAGTCAACACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCA
TGCCCT**TAA**ACTGGCATCCGGCCTTGTGGGAGAATAATGTCGCCGTTGTCACATCAGCTGAC
ATGACCTGGAGGGGTTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTTGTACTG
GGATTTGTGAATAAACTTGATACACCA

FIGURE 140

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66675

><subunit 1 of 1, 247 aa, 1 stop

><MW: 25335, pI: 7.00, NX(S/T): 0

MHLARLVGSCSLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITHAG
REVEKVFNGLSNMGSHTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHGVNNAAG
QAGKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQN
AHNGVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWRSVANIMP

Important features of the protein:

Signal peptide:

amino acids 1-25

Homologous region to circumsporozoite (CS) repeats:

amino acids 35-225

FIGURE 141

CTCCGGGTCCCCAGGGGCTGCGCCGGGCGGGCCTGGCAAGGGGGACGAGTCAGTGGACACTCCAGGAAGAGCGGC
 CCCGCGGGGGGCGATGACCGTGCGCTGACCCTGACTCACTCCAGGTCCGGAGGCGGGGGCCCCCGGGGCGACTCG
 GGGGCGGACCGCGGGGCGGAGCTGCCGCCCTGAGTCCGGCCGAGCCACCTGAGCCCGAGCCGCGGGACACCGTC
 GCTCCTGCTCTCCGAATGCTGCGCACCCGCGATGGGCTGAGGAGCTGGCTCGCCGCCCATGGGGCGCGTGGCCG
 CCTCGGCCACCGCTGCTGCTGCTCCTGCTGCTGCTGCTCCTGCTGCAGCCGCCGCTCCGACCTGGGCGCTCAGC
 CCCCAGATCAGCCTGCCTCTGGGCTCTGAAGAGCGGCCATTCCTCAGATTGGAAGCTGAACACATCTCCAACTAC
 ACAGCCCTTCTGCTGAGCAGGGATGGCAGGACCCGTGTACGTGGGTGCTCGAGAGGCCCTCTTTGCACTCAGTAGC
 AACCTCAGCTTCTGCCAGGCGGGGAGTACCAGGAGCTGCTTTGGGGTGACAGCGCAGAGAAGAAACAGCAGTGC
 AGCTTCAAGGGCAAGGACCCACAGCGCGACTGTCAAACTACATCAAGATCCTCCTGCCGCTCAGCGGCAGTCAC
 CTGTTACCTGTGGCACAGCAGCCTTACGCCCCATGTGTACCTACATCAACATGGAGAACTTACCCTGGCAAGG
 GACGAGAAGGGGAATGTCTCCTGGAAGATGGCAAGGGCCGTTGTCCCTTCGACCCGAATTTCAAGTCCACTGCC
 CTGGTGGTTGATGGCGAGCTCTACACTGGAACAGTCAGCAGCTTCCAAGGGAATGACCCGGCCATCTCGCGGAGC
 CAAAGCCTTCGCCCCACCAAGACCGAGAGCTCCCTCAACTGGCTGCAAGACCCAGCTTTTGTGGCCTCAGCCTAC
 ATTCTGAGAGCCTGGGCAGCTTGAAGGCGATGATGACAAGATCTACTTTTTTCTCAGCGAGACTGGCCAGGAA
 TTTGAGTTCTTTGAGAACACCATTGTGTCCCGCATTTGCCCGCATCTGCAAGGGCGATGAGGGTGAGAGCGGGTG
 CTACAGCAGCGCTGGACCTCCTTCTCAAGGCCAGCTGCTGTGCTCACGGCCCGACGATGGCTTCCCCTTCAAC
 GTGCTGCAGGATGTCTTACGCTGAGCCCCAGCCCCAGGACTGGCGTGACACCCCTTTTCTATGGGGTCTTCACT
 TCCCAGTGGCACAGGGGAACACAGAAGGCTCTGCCGTCTGTCTTCAACAATGAAGGATGTGCAGAGAGTCTTC
 AGCGGCCTCTACAAGGAGGTGAACCGTGAGACACAGCAGTGGTACACCGTGACCCACCCGGTGCCACACCCCGG
 CCTGGAGCGTGATCACCAACAGTGGCCGGGAAGGAAGATCAACTCATCCCTGCAGCTCCCAGACCGCGTGCTG
 AACTTCTCAAGGACCACTTCTGATGGACGGGCGAGGTCCGAAGCCGATGCTGCTGCTGCAGCCCCAGGCTCGC
 TACCAGCGCGTGGCTGTACACCGCGTCCCTGGCCTGCACCACACCTACGATGTCTCTTCTGGGCACTGGTGAC
 GGCCGGCTCCACAAGGCAGTGAGCGTGGGCCCCCGGTGCACATCATTGAGGAGCTGCAGATCTTCTCATCGGGA
 CAGCCCGTGCAAGATCTGCTCCTGGACACCCACAGGGGGCTGCTGTATGCGGCCCTCACACTCGGGCGTAGTCCAG
 GTGCCCATGGCCAACCTGCAGCCTGTACCGGAGCTGTGGGGACTGCCCTCCTCGCCCGGGACCCCTACTGTGCTTGG
 AGCGGCTCCAGCTGCAAGCACGTGAGCCTCTACCAGCCTCAGCTGGCCACCAGGCCGTGGATCCAGGACATCGAG
 GGAGCCAGCGCCAAGGACCTTTGAGCGCGCTTTCGGTTGTGTCCCGCTCTTTGTACCAACAGGGGAGAAAGCCA
 TGTGAGCAAGTCCAGTTCAGCCCAACACAGTGAACACTTTGGCCTGCCCGCTCCTCTCCAACCTGGCGACCCGA
 CTCTGGCTACGCAACGGGGCCCCCGTCAATGCCCTCGGCCCTCCTGCCACGTGCTACCCACTGGGGACCTGCTGCTG
 GTGGGCACCCAAACAGCTGGGGGAGTTCCAGTGTGGTCACTAGAGGAGGGCTTCCAGCAGCTGGTAGCCAGCTAC
 TCCCCAGAGGTGGTGGAGGACGGGGTGGCAGACCAACAGATGAGGGTGGCAGTGTACCCGTATTATCAGCACA
 TCGCGTGTGAGTGACACAGCTGGTGGCAAGGCCAGCTGGGGTGACAGACAGGTCTTACTGGAAGGAGTTCTGGTG
 ATGTGCACGCTCTTTGTGCTGGCCGTGCTGCTCCAGTTTTATTCTTGCTCTACCGGCACCGGAACAGCATGAAA
 GTCTTCTGAAGCAGGGGGAATGTGCCAGCGTGACCCCAAGACCTGCCCTGTGGTGTGCCCCCTGAGACCCGC
 CCACTCAACGGCCTAGGGCCCCCTAGCACCCCGCTCGATCACCGAGGGTACCAGTCCCTGTGAGACAGCCCCCG
 GGGGCCCGAGTCTTCACTGAGTCAGAGAAGAGGCCACTCAGCATCCAAGACAGCTTCGTGGAGGTATCCCCAGTG
 TGCCCCCGGGCCCCGGGTCCGCCCTTGGCTCGGAGATCCGTGACTCTGTGGTGTGAGAGCTGACTTCCAGAGGACGC
 TGCCCTGGCTTCAGGGGCTGTGAATGCTCGGAGAGGGTCAACTGGACCTCCCTCCGCTCTGCTCTTCTGGAAC
 ACGACCGTGGTGCCCGGCCCTTGGGAGCCTTGGAGCCAGCTGGCCTGCTGCTCTCCAGTCAAGTAGCGAAGCTCC
 TACCACCCAGACACCCAAACAGCCGTGGCCCCAGAGGTCTGGCCAAATATGGGGGCTGCCTAGGTTGGTGGA
 CAGTGCTCCTTATGTAACTGAGCCCTTGTGTTAAAAACAATTCCAAATGTGAAACTAGAATGAGAGGGAAGAG
 ATAGCATGGCATGCAGCACACGGCTGCTCCAGTTTATGGCCTCCAGGGGTGCTGGGGATGCATCCAAAGTGG
 TTGTCTGAGACAGAGTTGGAACCCCTACCAACTGGCCTTTCACCTTCCACATTATCCCGCTGCCACCGGCTGC
 CCTGTCTCACTGCAGATTACAGGACAGCTTGGGCTGCGTGCGTTCTGCCTTGCCAGTCAGCCGAGGATGTAGTTG
 TTGCTGCCGTGCTCCACCACCTCAGGGACCAGAGGGCTAGGTTGGCACTGCGGCCCTCACCAGGTCTGGGCTC
 GGACCCAACTCCTGGACCTTTCCAGCCTGTATCAGGCTGTGGCCACACGAGAGGACAGCGGAGCTCAGGAGAGA
 TTTCTGTGACAATGTACGCCCTTCCCTCAGAATTCAGGGAAGAGACTGTGCGCTGCCCTTCTCCGTTGTTGCGTGA
 GAACCCGTGTGCCCCCTCCACCATATCCACCCTCGCTCCATCTTTGAACTCAAACACGAGGAACCTAACTGCACC
 CTGGTCTCTCCCGAGTCCCGAGTTACCCCTCCATCCCTCACCTTCTCCACTCTAAGGGATATCAACACTGCCC
 AGCACAGGGGCCCTGAATTTATGTGGTTTTTATACATTTTTTAATAAGATGCACTTTATGTATTTTTTAATAAA
 GTCTGAAGAATTACTGTTTTAAAAA

FIGURE 142

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA67962

><subunit 1 of 1, 837 aa, 1 stop

><MW: 92750, pI: 7.04, NX(S/T): 6

MLRTAMGLRSWLAAPWGALPPRPPLLLLLLLLLLLLLLQPPPPTWALSPRISLPLGSEERPFLRF
 EAEHISNYTALLLSRDGRTLYVGAREALFALSSNLSFLPGGEYQELLWGADAEKKQQCSFKG
 KDPQRDCQNYIKILLPLSGSHLFTCGTAAAFSPMCTYINMENFTLARDEKGNVLLEDGKGRCP
 FDPNFKSTALVVDGELYTGTVSSSQGNDPAISRSQSLRPTKTESSLNWLQDPAFVASAYIPE
 SLGSLQGDDDKIYFFFESETGQEFFEFENTIVSRIARICKGDEGGERVLQQRWTSFLKAQLLC
 SRPDDGFPPFNVLQDVFTLSPSPQDWRDTLFYGVFTSQWHRGTTEGSAVCVFTMKDVQRVFSG
 LYKEVNRETQQWYTVTHPVPTPRPGACITNSARERKINSSQLPDRVLNFKDHFLLMDGQVR
 SRMLLLQPQARYQRVAVHRVPGLHHTYDVLFLGTGDGRLHKAVSVGPRVHIIEELQIFSSGQ
 PVQNLLLDTHRGLLYAASHSGVVQVPMANCSLYRSCGDCLLARDPYCAWSGSSCKHVSQYQP
 QLATRPWIQDIEGASAKDLCSASSVSPSFVPTGKPCQVQFQPNVTNLTACPLLSNLATR
 LWLRNGAPVNASASCHVLPTGDLVLVGTQQLGEFQCWSLEEGFQQLVASYCPEVVEDGVADQ
 TDEGGSPVPIISTSRVSAPAGGKASWGADRSYWKEFLVMCTLFVLAVLLPVLFLLYRHRNSM
 KVFLKQGECAVHPKTCPPVLPPEPTRPLNGLGPPSTPLDHRGYQSLSDSPPGARVFTSEKR
 PLSIQDSFVEVSPVCPRPRVRLGSEIRDSVV

Transmembrane domains:

amino acids 23-46 (type II), 718-738

FIGURE 143A

FIGURE 143B

TTAATTTTATATTCCTTACTGTTTTACTAAATATTAAGTGTTCTTTGACAATTTTGGTGCTCATGTGTTTTGGG
GACAAAAGTGAAATGAATCTGTCATTATACCAGAAAGTTAAATTCTCAGATCAAATGTGCCTTAATAAATTTGTT
TTCATTTAGATTTCAAACAGTGATAGACTTGCCATTTTAATACACGTCATTGGAGGGCTGCGTATTTGTAAATAG
CCTGATGCTCATTGGGAAAAATAAACAGTGAACAATATTTTCTATTGTACTTTTCGAACCATTTTGTCTCATT
ATTCCTGTTTTAGCTGAAGAATTGTATTACATTTGGAGAGTAAAAACTTAAACACGAAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 144

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68836

><subunit 1 of 1, 802 aa, 1 stop

><MW: 91812, pI: 9.52, NX(S/T): 3

MAARGRRAWLSVLLGLVLGFLASRLVLPRASELKRAGPRRRASPEGCRSGQAAASQAGGAR
GDARGAQLWPPGSDPDGGPRDRNFLFVGVMQAQKYLQTRAVAAAYRTWSKTI PGKVQFFSSEG
SDTSVPIPVVPLRGVDDSYPPQKKSFMMLKYMHDHYLDKYEFMRADDDVYIKGDRLENFLR
SLNSSEPLFLGQTGLGTTEEMGKLALEPGENFCMGGPGVIMSREVLRRMVPHIGHKCLREMYT
THEDVEVGRCVRRFAGVQCWVSYEMRQLFYENYEQNKKGYIRDLHNSKIHQAITLHPNKNPP
YQYRLHSYMLSRKISELRHRTIQLHREIVLMSKYSNTEIHKEDLQLGI PPSFMRFQPRQREE
ILEWEFLTGKYLYSAVDGQPPRRGMDSAQREALDDIVMQVMEMINANAKTRGRIIDFKEIQY
GYRRVNPMYGAEYILDLLLLYKKHKGKKMTVPVRRHAYLQQTFSKIQFVEHEELDAQELAKR
INQESGSLSFSLNSLKKLVFPQLPGSKSEHKEPKDKKINILIPLSGRFDMFVRFMGNFEKTC
LIPNQNVKLVVLLFNSDSNPDKAKQVELMRDYRIKYPKADMQILPVSGEFSRALALEVGSSQ
FNNESELLFFCDVDLVFTTEFLQRCRANTVLGQQIYFPIIFSQYDPKIVYSGKVPSDNHFAFT
QKTGFWRNYGFGITCIYKGD LVRVGGFDVSIQGWGLEDDVLFNKVVQAGLKTFRSQEVGVVH
VHHPVFCDPNLDPKQYKMCLGSKASTYGSTQQLAEMWLEKNDPSYSKSSNNNGSVRTA

Signal peptide:

amino acids 1-23

FIGURE 145

GGACAACCGTTGCTGGGTGTCCCAGGGCCTGAGGCAGGACGGTACTCCGCTGACACCTTCCC
TTTCGGCCTTGAGGTTCCCAGCCTGGTGGCCCCAGGACGTTCCGGTCGCATGGCAGAGTGCT
ACGGACGACGCCT**ATG**AAGCCCTTAGTCCTTCTAGTTGCGCTTTTGCTATGGCCTTCGTCTG
TGCCGGCTTATCCGAGCATAACTGTGACACCTGATGAAGAGCAAACTTGAATCATTATATA
CAAGTTTTAGAGAACCTAGTACGAAGTGTTCCCTCTGGGGAGCCAGGTCGTGAGAAAAAATC
TAACTCTCCAAAACATGTTTATTCTATAGCATCAAAGGGATCAAAATTTAAGGAGCTAGTTA
CACATGGAGACGCTTCAACTGAGAATGATGTTTTAACCAATCCTATCAGTGAAGAACTACA
ACTTTCCTTACAGGAGGCTTCACACCGGAAATAGGAAAGAAAAACACACGGAAAGTACCCC
ATTCTGGTCGATCAAACCAAACAATGTTTCCATTGTTTTGCATGCAGAGGAACCTTATATTG
AAAATGAAGAGCCAGAGCCAGAGCCGGAGCCAGCTGCAAAACAACTGAGGCACCAAGAATG
TTGCCAGTTGTTACTGAATCATCTACAAGTCCATATGTTACCTCATAACAAGTCACCTGTCAC
CACTTTAGATAAGAGCACTGGCATTGAGATCTCTACAGAATCAGAAGATGTTCCCTCAGCTCT
CAGGTGAAACTGCGATAGAAAAACCCGAAGAGTTTGGAAGCACCCAGAGAGTTGGAATAAT
GATGACATTTTGAAAAAAATTTTAGATATTAATTCACAAGTGCAACAGGCACTTCTTAGTGA
CACCAGCAACCCAGCATATAGAGAAGATATTGAAGCCTCTAAAGATCACCTAAAACGAAGCC
TTGCTCTAGCAGCAGCAGCAGAACATAAATTAAAAACAATGTATAAGTCCCAGTTATTGCCA
GTAGGACGAACAAGTAATAAAATTGATGACATCGAACTGTTATTAACATGCTGTGTAATTC
TAGATCTAAACTCTATGAATATTTAGATATTAAATGTGTTCCACCAGAGATGAGAGAAAAAG
CTGCTACAGTATTCAATACATTAAAAAATATGTGTAGATCAAGGAGAGTCACAGCCTTATTA
AAAGTTTAT**TAA**ACAATAATATAAAAAATTTTAAACCTACTTGATATTCCATAACAAAGCTGA
TTTAAGCAAACTGCATTTTTTTCACAGGAGAAATAATCATATTTCGTAATTTCAAAAGTTGTAT
AAAAATATTTTCTATTGTAGTTCAAATGTGCCAACATCTTTATGTGTCATGTGTTATGAACA
ATTTTCATATGCACTAAAAACCTAATTTAAATAAAATTTTGGTTCAGGAAAAA

FIGURE 146

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68864
```

```
><subunit 1 of 1, 350 aa, 1 stop
```

><MW: 39003, pI: 5.59, NX(S/T): 1

MKPLVLLVALLLWPSSVPAYPSITVTPDEEQNLNHYIQVLENLVRVSPSGEPGREKKSNSPK
 HVYSIASKGSKFKELVTHGDASTENDVLTNPISEETTTFTPTGGFTPEIGKKKHTESTPFWSI
 KPNNVSIVLHAEOPYIENEEPEPEPEPAAKQTEAPRMLPVVTESSTSPYVTSYKSPVTTLDK
 STGIEISTESEDVPLSGETAIEKPEEFGKHPESWNNDDILKKILDINSQVQQALLSDTSNP
 AYREDIEASKDHLKRSALAAAAEHKLKTMYSQLLPVGRTSNKIDDIETVINMLCNSRSKL
 YEYLDIKCVPPEMREKAATVFNTLKNMCRSRRVTALLKVY

Signal peptide:

amino acids 1-19

FIGURE 147

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTG
ACCTCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTCATAGTGTGAGATCAACCCACA
GGAATATCC**ATG**GGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACA
GTGGCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCT
CCTGCTCCCTCTTTCCTGAGACCAGTGCAGAGGCCTATGGAAGTGC GGTTCTTCAGGAATCAG
TTCCATGCTGTGGTCCACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACA
GTATCGAGGGAGAACTGAGTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGC
TAAAAAACATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTCAGTTCCCAGATTTAC
GATGAGGAGGCCACCTGGGAGCTGCGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCAT
CGTGGGATATGTTGACGGAGGTATCCAGTTACTCTGCCTGTCTCAGGCTGGTTCCTCCAGC
CCACAGCCAAGTGGAAAGGTCCACAAGGACAGGATTTGTCTTCAGACTCCAGAGCAAATGCA
GATGGGTACAGCCTGTATGATGTGGAGATCTCCATTATAGTCCAGGAAAATGCTGGGAGCAT
ATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAG
AGACGTTTTTCCAGCCCTCACCTTGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGT
GCCCTGTGTGGTGTGTGTCATGGGGATGATAATTGTTTTCTTCAAATCCAAAGGGAAAATCCA
GGCGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACG
CAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAA
ACTGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAA
GAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGGAC
AAAATGTAGGGTGGTATGTGGGAGTGTGTGCGGATGACGTAGACAGGGGGAAGAACAATGTG
ACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTTAC
ATTCAATCCCCATTTTATCAGCCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCC
TGGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACC
CTGCTGACATGTCAGTTTGAAGGCTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGA
GGAAAAGGGGACTCCCATATTATATGTCCAGTGTCTGGGGAT**TGA**GACAGAGAAGACCCTG
CTTAAAGGGCCCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGC
CCCAGCTTCCTCTCCGGAGCCTGCGCACAGAGAGTCACGCCCCCCTCTCCTTTAGGGAGC
TGAGGTTCCTCTGCCCTGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGAT
TGGCCTGACCTGTGGGAGTCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCA
CATTAGGTTTGTGTTTGTGAAAACCTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCC
CAGGCTCCTCATTTGTAGTACACGGACAGTGATTCCCTGCCTCACAGGTGAAGATTAAAGAGA
CAACGAATGTGAATCATGCTTGCAGGTTTGAAGGACAGTGTTTGTCTAATGATGTGTTTTTA
TATTATACATTTTCCCACCATAAACTCTGTTTGCTTATTCCACATTAATTTACTTTTCTCTA
TACCAAATCACCCATGGAATAGTTATTGAACACCTGCTTTGTGAGGCTCAAAGAATAAAGAG
GAGGTAGGATTTTCACTGATTCTATAAGCCCAGCATTACCTGATACCAAAAACAGGCAAAG
AAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCCCTCATTAACACAGACACAAAAA
TTCTAAATAAAATTTTAAACAAATTAACTAAACAATATATTTAAAGATGATATATAACTACT
CAGTGTGGTTTGTCCCACAAATGCAGAGTTGGTTTAAATATTTAAATATCAACCAGTGTAATT
CAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAAAAAAAAAAAAAA

FIGURE 148

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68866
```

><subunit 1 of 1, 466 aa, 1 stop

><MW: 52279, pI: 6.16, NX(S/T): 2

MAFVLLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHA
VVHLYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEE
ATWELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGY
SLYDVEISIIVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALC
GVVMGMIIVFFKSKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVT
HRKAPQEVPHSEKRFTRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLS
PNNGYWVLRLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLTT
CQFEGLLRPYIQHAMYDEEKGTPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 149

CCTTCACAGGACTCTTCATTGCTGGTTGGCA**ATG**ATGTATCGGCCAGATGTGGTGAGGGCTA
 GGAAAAGAGTTTGTGGGAACCCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTC
 CTGGCAGTGTGCATTGGACTCACTGTTTATTATGTGAGATATAATCAAAAGAAGACCTACAA
 TTACTATAGCACATTGTCAATTTACAACCTGACAACTATATGCTGAGTTTGGCAGAGAGGCTT
 CTAACAATTTTACAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAA
 TCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGG
 AGTGTGGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAG
 ATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTA
 GATCCTCACTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCA
 TTGCTGCGGAACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAG
 AAGTAGAAGAGGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGT
 GGAGCAACCTTAATTAATGCCACATGGCTTGTGAGTGCTGCTCACTGTTTTACAACATATAA
 GAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACAATAAAACCTTCGAAAATGAAACGGG
 GTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACATGACTATGATATTTCT
 CTTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTCCCTGA
 TGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAA
 ATGATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAAC
 TGCAATGAACCTCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTT
 AGAAGGAAAAACAGATGCATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTA
 GAGATATCTGGTACCTTGCTGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCCAACAAG
 CCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTGGATTACTTCAAAAACCTGGTAT**CTA**
AGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTTGTTTTTTGGGTGTGGAGGCCATT
 TTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAAC
 TGTTTGCTTGATGCATGTATTTTCTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTG
 CCAGATCAACTCTGTCACTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATA
 ATACAATATTACATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTTGTCAGAATTTTGAC
 TTGTTGACATAAATTTGTAATGCATATATAACAATTTGAAGCACTCCTTTTCTTCAGTTCCTC
 AGCTCCTCTCATTTTCAGCAAATATCCATTTTCAAGGTGCAGAACAAGGAGTGAAAGAAAATA
 TAAGAAGAAAAAATCCCCTACATTTTATTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGT
 GGAATATTAGAAATGATCATATTCATTATGAAAGGTCAAGCAAAGACAGCAGAATACCAATC
 ACTTCATCATTTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAAGCCAAGATATA
 TCCTTATTTTTCATTTCCAAACAACCTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTG
 ACCTATAATAATTATACAACTTCATGCAATGTACTTGTTCTAAGCAAATTAAGCAAATAT
 TTATTTAACATTGTTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 150

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68871

><subunit 1 of 1, 423 aa, 1 stop

><MW: 47696, pI: 8.96, NX(S/T): 3

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYSTLSFTTD
KLYAEFGREASNNFTEMSQRLESMVKNAFYKSPREEFVKSQVIKFSQQKHGVLAHMLLICR
FHSTEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTL
GQSLRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFG
VTIKPSKMKRGLRRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDV
MFVTGFGALKNDGYSONHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGD
SGGPLVSSDARDIWYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 151

GTCGAAGGTTATAAAAGCTTCCAGCCAAACGGCATTGAAGTTGAAGATACAACCTGACAGCA
CAGCCTGAGATCTTGGGGATCCCTCAGCCTAACACCCACAGACGTCAGCTGGTGGATTCCCC
CTGCATCAAGGCCTACCCACTGTCTCC**ATG**CTGGGCTCTCCCTGCCTTCTGTGGCTCCTGGC
CGTGACCTTCTTGGTTCCCAGAGCTCAGCCCTTGGCCCCCTCAAGACTTTGAAGAAGAGGAGG
CAGATGAGACTGAGACGGCGTGGCCGCCTTTGCCGGCTGTCCCCTGCGACTACGACCACTGC
CGACACCTGCAGGTGCCCTGCAAGGAGCTACAGAGGGTCGGGCCGGCGGCCTGCCTGTGCCC
AGGACTCTCCAGCCCCGCCAGCCGCCGACCCGCCGCGCATGGGAGAAGTGCGCATTGCGG
CCGAAGAGGGCCGCGCAGTGGTCCACTGGTGTGCCCCCTTCTCCCCGGTCTCTCCACTACTGG
CTGCTGCTTTGGGACGGCAGCGAGGCTGCGCAGAAGGGGGCCCCCGCTGAACGCTACGGTCCG
CAGAGCCGAACCTGAAGGGGGCTGAAGCCAGGGGGGCATTTATGTCGTTTGCGTAGTGGCCGCTA
ACGAGGCCGGGGCAAGCCGCGTGCCCCAGGCTGGAGGAGAGGGCCTCGAGGGGGCCGACATC
CCTGCCTTCGGGCCTTGACAGCCGCCTTGCGGTGCCGCCCAACCCCCGCACTCTGGTCCACGC
GGCCGTCGGGGTGGGCACGGCCCTGGCCCTGCTAAGCTGTGCCGCCCTGGTGTGGCACTTCT
GCCTGCGCGATCGCTGGGGCTGCCCCGCGCCGAGCCGCCGCCGAGCCGCAGGGGCGCTCT**TGA**
AAGGGGCCTGGGGGCATCTCGGGCACAGACAGCCCCACCTGGGGCGCTCAGCCTGGCCCCCG
GGAAAGAGGAAAACCCGCTGCCTCCAGGGAGGGCTGGACGGCGAGCTGGGAGCCAGCCCCAG
GCTCCAGGGCCACGGCGGAGTCATGGTTCTCAGGACTGAGCGCTTGTTTAGGTCCGGTACTT
GGCGCTTTGTTTCCTGGCTGAGGTCTGGGAAGGAATAGAAAGGGGCCCCCAATTTTTTTTTTA
AGCGGCCAGATAATAAATAATGTAACCTTTGCGGTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 152

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68874

><subunit 1 of 1, 238 aa, 1 stop

><MW: 25262, pI: 6.44, NX(S/T): 1

MLGSPCLLWLLAVTFLVPRAQPLAPQDFEEEEADETETAWPPLPAVPCDYDHCRLQVPCKE
LQRVGPAACLCPGLSSPAQPPDPGRMGEVRIAAEEGRAVVHWCAPFSPVLHYWLLLWDGSEA
AQKGPPLNATVRRRAELKGLKPGGIYVVCVVAANEAGASRVPQAGGEGLEGADIPAFGPC SRL
AVPPNPRTL VHAAVGVGTALALLSCAALVWHFCLRDWRWGCPRRAAARAAGAL

Important features of the protein:**Signal peptide:**

amino acids 1-20

Transmembrane domain:

amino acids 194-220

N-glycosylation site.

amino acids 132-135

FIGURE 153

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCC
 CTGCCCCGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGG
 CCCAGCGCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGC
 TTCTCCCTTACGGGGCTCACA**ATGG**CCAGAGAAGATTCGGTGAAGTGTCTGCGCTGCCTGCT
 CTACGCCCTCAATCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGA
 TGAGGGACTACCTAAATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTC
 ATTTTGACTTACTTTTCTGTGGTTCATCCGGTCATGATTGCTGTTTGCTGTTTCTTATCAT
 TGTGGGGATGTTAGGATATTGTGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCAATGGTACT
 TTGGAAGTTTGCTTGTCATTTTCTGTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAG
 GAACTTATGGTTCAGTACAATGGTCAGATATGGTCACTTTGAAAGCCAGGATGACAAATTA
 TGGATTACCTAGATATCGGTGGCTTACTCATGCTTGGAATTTTTTTTCAGAGAGAGTTTAAGT
 GCTGTGGAGTAGTATATTTCACTGACTGGTTGGAAATGACAGAGATGGACTGGCCCCCAGAT
 TCCTGCTGTGTTAGAGAATTCCCAGGATGTTCCAAACAGGCCCCACCAGGAAGATCTCAGTGA
 CCTTTATCAAGAGGGTGTGGGAAGAAAATGTATTCTTTTTTGAGAGGAACCAAACTGC
 AGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAATCCTGGCCATGATTCTCACC
 ATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGGACAGACCAATGATGTC
 CTTGAAGAATGACAACTCTCAGCACCTGTCAATGTCCCTCAGTAGAACTGTTGAAACCAAGCC
 TGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACTTTGAGATGGAG
 GAGTTA**TAAAA**AGAAATGTCACAGAAGAAAACCACAACTTGTTTTATTGGACTTGTGAATT
 TTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAATAACA
 CCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTCAC
 CACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGC
 CTGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTT
 CGCATCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTT
 CTACCAACTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATA
 ACTTTTATTACTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTC
 AATATTGGTGACTACCTAAATGTGATTTTTTGCTGGTTACTAAAATATTCTTACCCTTAAAA
 GAGCAAGCTAACACATTGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTAA
 TCTGTATAATTGAGTCGATTTTCAGTTCTGATAATGTTAAGAATAACCATTTATGAAAAGGAAA
 ATTTGTCTGTATAGCATCATTATTTTTAGCCTTTCCTGTTAATAAAGCTTTACTATTCTGT
 CCTGGGCTTATATTACACATATAACTGTTATTTAAATACTTAACCACTAATTTTGAAAATTA
 CCAGTGTGATACATAGGAATCATTATTCAGAAATGTAGTCTGGTCTTTAGGAAGTATTAATA
 GAAAATTTGCACATAACTTAGTTGATTGAGAAAGGACTTGATGCTGTTTTTCTCCAAATG
 AAGACTCTTTTTGACACTAAACACTTTTTAAAAAGCTTATCTTTGCCTTCTCCAAACAAGAA
 GCAATAGTCTCCAAGTCAATATAAATTCTACAGAAAATAGTGTTCTTTTTCTCCAGAAAAT
 GCTTGTGAGAATCATTAAACATGTGACAATTTAGAGATTCTTTGTTTTATTTCACTGATTA
 ATATACTGTGGCAAATTACACAGATTATTAAATTTTTTTACAAGAGTATAGTATATTTATTT
 GAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTATTCTCAGAATATGGAA
 AGAAAATTAAATGTGTCAATAAATATTTTCTAGAGAGTAA

FIGURE 154

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68880

><subunit 1 of 1, 305 aa, 1 stop

><MW: 35383, pI: 5.99, NX(S/T): 0

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPV
VHPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQ
WSDMVTLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREF
PGCSKQAHQEDLSPLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWAL
YYDRREPGTDQMMSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 155

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCC
 TGCACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGA
 CCTGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAG
 ACCGGGAGGATCACAGAGCCAGC**ATG**TTACAGGATCCTGACAGTGATCAACCTCTGAACAGC
 CTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGAT
 CCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCCTCATCAAGG
 TGATTCTGGATAAATACTACTTCCCTGCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAG
 CTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTT
 CCCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGG
 ACTCGGCCACAGGGAAGTGGTTCTCTGCCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAG
 ACAGCCTGTAGGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGA
 TGTTGTTGAAATCACAGAAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTC
 TCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCTGTGGGAAGAGCCTGAAGACCCCCCGT
 GTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGGTCAGCATCCAGTACGA
 CAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCAGCCCACT
 GCTTCAGGAAACATAACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGC
 AGCTTCCCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAA
 AGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGC
 CCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGA
 TGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCA
 GGTCATTGACAGCACACGGTGCAATGCAGACGATGCGTACCAGGGGGGAAGTCACCGAGAAGA
 TGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCC
 CTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGG
 GGGCCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATG
 TCTGGAAGGCTGAGCTG**TAA**TGCTGCTGCCCCTTTGCAGTGCTGGGAGCCGCTTCCTTCCTG
 CCCTGCCCCACCTGGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCC
 CTCTGCCCCACAGCCTCAGCATTCTTGGAGCAGCAAAGGGCCTCAATTCCCTGTAAGAGACCC
 TCGCAGCCCAGAGGCGCCCAGAGGAAGTCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCC
 AGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAA
 GGAAGTTTCCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAGGCCAGATCACTGTGGG
 CTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTCACCCATCCCCAAGCCTA
 CTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTT
 ACCTACTGTTGTCAATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCT
 CTGGCAAAAAAAAAAAAA

FIGURE 156

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68885

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47644, pI: 5.18, NX(S/T): 2

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVVLIVLIKVILDKYYF
LCGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWF
SACFDNFTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSL
HCLACGKSLKTPRVVGEEASVDSWPWQVSIQYDKQHVC GGSILDPHWVLTAAHCFRKHTDV
FNWKVRAGSDKLGSEFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFD
EELTPATPLWIIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAPIPE
GGVDTQCQGDSSGGLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 157

GGGCTGAGGCACTGAGAGACCGGAAAGCCTGGCATTCCAGAGGGAGGGAAACGCAGCGGCATCCCCAGGCTCCAG
 AGCTCCCTGGTGACAGTCTGTGGCTGAGC**ATG**GCCCTCCCAGCCCTGGGCCTGGACCCCTGGAGCCTCCTGGGCC
 TTTTCCTCTTCCAACCTGCTTCAGCTGCTGCTGCCGACGACGACCGCGGGGGAGGCGGGCAGGGGGCCCATGCCCA
 GGGTCAGATACTATGCAGGGGATGAACGTAGGGCACTTAGCTTCTTCCACCAGAAGGGCCTCCAGGATTTTGACA
 CTCTGCTCCTGAGTGGTGATGGAAATACTCTCTACGTGGGGGCTCGAGAAGCCATTCTGGCCTTGGATATCCAGG
 ATCCAGGGGTCCCCAGGCTAAAGAACATGATACCGTGGCCAGCCAGTGACAGAAAAAGAGTGAATGTGCCTTTA
 AGAAGAAGAGCAATGAGACACAGTGTTCAACTTCATCCGTGTCTGGTTTCTTACAATGTACCCATCTCTACA
 CCTGCGGCACCTTCGCCTTCAGCCCTGCTTGTACCTTCATTGAACCTCAAGATTCTTACCTGTTGCCCATCTCGG
 AGGACAAGGTCATGGAGGGAAAAGGCCAAAGCCCCCTTTGACCCCGCTCACAAGCATACGGCTGTCTGGTGATG
 GGATGCTCTATTCTGGTACTATGAACAACCTTCTGGGCAGTGAGCCCATCCTGATGCGCACACTGGGATCCCAGC
 CTGTCTCAAGACCGACAACCTTCTCCGCTGGCTGCATCATGACGCCCTCCTTTGTGGCAGCCATCCCTTCGACCC
 AGGTGCTCTACTTCTTCTTCGAGGAGACAGCCAGCGAGTTTGACTTCTTTGAGAGGCTCCACACATCGCGGGTGG
 CTAGAGTCTGCAAGAATGACGTGGGCGGCGAAAAGCTGCTGCAGAAGAAGTGGACCACCTTCCTGAAGGCCCAGC
 TGCTCTGACCCAGCCGGGGCAGCTGCCCTTCAACGTTCATCCGCCACGCGGTCTGCTCCCCGCCGATTCTCCCA
 CAGCTCCCCACATCTACGCAGTCTTCACCTCCAGTGGCAGGTTGGCGGGACCAGGAGCTCTGCGGTTTGTGCCT
 TCTCTCTCTTGGACATTGAACGTGTCTTTAAGGGGAAATACAAAGAGTTGAACAAAGAACTTCACGCTGGACTA
 CTTATAGGGGCCCTGAGACCAACCCCCGGCCAGGCAGTTGCTCAGTGGGCCCCCTCCTCTGATAAGGCCCTGACCT
 TCATGAAGGACCATTTCTGATGGATGAGCAAGTGGTGGGGACGCCCTGCTGGTGAAATCTGGCGTGGAGTATA
 CACGGCTTGCACTGGAGACAGCCAGGGCCTTGATGGGCACAGCCATCTTGTCTGTACCTGGGAACCACCACAG
 GGTGCTCCACAAGGCTGTGGTAAGTGGGGACAGCAGTGCTCATCTGGTGGAAGAGATTAGCTGTTCCCTGACC
 CTGAACCTGTTGCAACCTGCAGCTGGCCCCCACCAGGGTGCAGTGTGTTGTAGGCTTCTCAGGAGGTGTCTGGA
 GGGTGCCCCGAGCCAACCTGTAGTGTCTATGAGAGCTGTGTGGACTGTGTCCTTGCCCCGGACCCCCACTGTGCCT
 GGGACCCTGAGTCCCGAACCTGTTGCCTCCTGTCTGCCCCAACCTGAACTCCTGGAAGCAGGACATGGAGCGGG
 GGAACCCAGAGTGGGCATGTGCCAGTGGCCCCATGAGCAGGAGCCTTCGGCCTCAGAGCCGCCCGCAAATCATT
 AAGAAGTCCTGGCTGTCCCCAACTCCATCCTGGAGCTCCCCGCCCCACCTGTCAGCCTTGGCCTCTTATTATT
 GGAGTCATGGCCCAGCAGCAGTCCCAGAAGCCTCTTCCACTGTCTACAATGGCTCCCTCTTGCTGATAGTGCAGG
 ATGGAGTTGGGGGTCTCTACCAGTGTGGGCAACTGAGAATGGCTTTTCATACCCTGTGATCTCCTACTGGGTGG
 ACAGCCAGGACCAGACCCTGGCCCTGGATCCTGAACTGGCAGGCATCCCCGGGAGCATGTGAAGGTCCCGTTGA
 CCAGGGTCAGTGGTGGGGCCGCCCTGGCTGCCCAGCAGTCTTACTGGCCCCACTTTGTCACTGTCACTGTCTCT
 TTGCCTTAGTGCTTTTCCAGGAGCCCTCATCATCTCGTGGCCTCCCCATTGAGAGCACTCCGGGCTCGGGGCAAGG
 TTCAGGGCTGTGAGACCCTGCGCCCTGGGGAGAAGGCCCGTTAAGCAGAGAGCAACACCTCCAGTCTCCCAAGG
 AATGCAGGACCTCTGCCAGTGATGTGGACGCTGACAACAACCTGCCTAGGCACTGAGGTAGCT**TAA**ACTCTAGGCA
 CAGGCCGGGGCTGCGGTGCAGGCACCTGGCCATGCTGGCTGGGCGGCCCAAGCACAGCCCTGACTAGGATGACAG
 CAGCACAAAAGACCACCTTTCTCCCCCTGAGAGGAGCTTCTGCTACTCTGCATCACTGATGACACTCAGCAGGGTG
 ATGCACAGCAGTCTGCCTCCCCCTATGGGACTCCCTTCTACCAAGCACATGAGCTCTCTAACAGGGTGGGGGCTAC
 CCCCAGACCTGCTCCTACACTGATATTGAAGAACCTGGAGAGGATCCTTCAGTTCTGGCCATTCCAGGGACCCTC
 CAGAAACACAGTGTTCAGAGACCCTAAAAACCTGCCTGTCCCAGGACCCTATGGTAATGAACACCAAACATC
 TAAACAATCATATGCTAACATGCCACTCCTGGAACTCCACTCTGAAGCTGCCGCTTTGGACACCAACACTCCCT
 TCTCCCAGGGTCATGCAGGATCTGCTCCCTCCTGCTTCCCTTACCAGTCGTGCACCGCTGACTCCCAGGAAGTC
 TTTCTGAAGTCTGACCACCTTTCTTCTTGCTTCAGTTGGGGCAGACTCTGATCCCTTCTGCCCTGGCAGAATGG
 CAGGGGTAATCTGAGCCTTCTTCACTCCTTTACCCTAGCTGACCCCTTCACCTCTCCCCCTCCCTTTTCTTTGT
 TTTGGGATTCAGAAAACCTGCTTGTGAGAGCTGTTTATTTTTTATTAAAAATATAAGGCTTAAAAAAA

FIGURE 158

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71166

><subunit 1 of 1, 761 aa, 1 stop

><MW: 83574, pI: 6.78, NX(S/T): 4

MALPALGLDPWSLLGLFLFQLLQLLLPTTTAGGGGQGMPRVRYAGDERRALSFFHQKGLQ
DFDTLLLSGDGNTLYVGAREAILALDIQDPGVPRLNMIWPASDRKKSECAFKKKSNETQC
FNFIRVLVSYNVTHLYTCGTFAFSPACTFIELQDSYLLPISEDKVMGKQSPFDPAHKHTA
VLVDGMLYSGTMNNFLGSEPILMRTLGSQPVLKTDNFLRWLHHDAFVAAIPSTQVVYFFFE
ETASEFDFFERLHTSRVARVCKNDVGGEKLLQKKWTTFLKAQLLCTQPGQLPFNVIRHAVLL
PADSPTAPHIYAVFTSQWQVGGTRSSAVCAFSLLDIERVFKGKYKELNKETSRWTTYRGPET
NPRPGSCSVGPSSDKALTFMKDHFLMDEQVVGTPLLVKSGVEYTRLAVETAQGLDGHSHLVM
YLGTTTGS LHKAVVSGDSSAHLVEEIQLFPDPEPVRNLQLAPTQGAVFVGFSGGVWRVPRAN
CSVYESCVDCVLARDPHCAWDPESTRCCLLSAPNLNSWKQDMERGNPEWACASGPMSRSLRP
QSRPQIIKEVLAVPNSILELPCPHLSALASYWWSHGPAAVPEASSTVYNGSLLLLIVQDGVGG
LYQCWATENGFSYPVISYWVDSQDQTLALDPELAGIPREHVKVPLTRVSGGAALAAQQSYWP
HFVTVTVL FALVLSGALIILVASPLRALRARGKVQGCETLRPGEKAPLSREQHLQSPKECRT
SASDVDADNNCLGTEVA

Signal peptide:

amino acids 1-30

Transmembrane domains:

amino acids 136-156, 222-247, 474-490, 685-704

FIGURE 159

AGGGTCCCTTAGCCGGGCGCAGGGCGCGCAGCCCAGGCTGAGATCCGCGGCTTCCGTAGAAG
 TGAGC**ATG**GCTGGGCAGCGAGTGCTTCTTCTAGTGGGCTTCCTTCTCCCTGGGGTCCTGCTC
 TCAGAGGCTGCCAAAATCCTGACAATATCTACAGTAGGTGGAAGCCATTATCTACTGATGGA
 CCGGGTTTCTCAGATTCTTCAAGATCACGGTCATAATGTCACCATGCTTAACCACAAAAGAG
 GTCCTTTTATGCCAGATTTTAAAAAGGAAGAAAAATCATATCAAGTTATCAGTTGGCTTGCA
 CCTGAAGATCATCAAAGAGAAATTTAAAAAGAGTTTTGATTTCTTTCTGGAAGAACTTTAGG
 TGGCAGAGGAAAATTTGAAAACCTATTAAATGTTCTAGAATACTTGGCGTTGCAGTGCAGTC
 ATTTTTTAAATAGAAAGGATATCATGGATTCCTTAAAGAATGAGAACTTCGACATGGTGATA
 GTTGAAACTTTTGACTACTGTCTTTCTGATTGCTGAGAAGCTTGGGAAGCCATTTGTGGC
 CATTCTTTCCACTTCATTCCGGCTCTTTGGAATTTGGGCTACCAATCCCCTTGTCTTATGTTT
 CAGTATTCGGTTCCTTGCTGACTGATCACATGGACTTCTGGGGCCGAGTGAAGAATTTTCTG
 ATGTTCTTTAGTTTCTGCAGGAGGCAACAGCACATGCAGTCTACATTTGACAACACCATCAA
 GGAACATTTACAGAAGGCTCTAGGCCAGTTTTGTCTCATCTTCTACTGAAAGCAGAGTTGT
 GGTTCATTAACCTGACTTTGCCTTTGATTTTGCTCGACCTCTGCTTCCCAACACTGTTTAT
 GTTGAGGCTTGATGGAAAAACCTATTAAACCAGTACCACAAGACTTGGAGAACTTCATTGC
 CAAGTTTGGGGACTCTGGTTTTGTCTTGTGACCTTGGGCTCCATGGTGAACACCTGTCAGA
 ATCCGGAAATCTTCAAGGAGATGAACAATGCCTTTGCTCACCTACCCCAAGGGGTGATATGG
 AAGTGTCAGTGTTCTCATTGGCCCAAAGATGTCCACCTGGCTGCAAATGTGAAAATTGTGGA
 CTGGCTTCCTCAGAGTGACCTCCTGGCTCACCCAAGCATCCGTCTGTTTGTACCCACGGCG
 GGCAGAAATAGCATAATGGAGGCCATCCAGCATGGTGTGCCCATGGTGGGGATCCCTCTCTTT
 GGAGACCAGCCTGAAAACATGGTCCGAGTAGAAGCCAAAAAGTTTGGTGTCTTCTATTAGTT
 AAAGAAGCTCAAGGCAGAGACATTGGCTCTTAAGATGAAACAAATCATGGAAGACAAGAGAT
 ACAAGTCCGCGGCAGTGGCTGCCAGTGTCTCCTGCGCTCCACCCGCTCAGCCCCACACAG
 CGGCTGGTGGGCTGGATTGACCACGTCTCCAGACAGGGGGCGCGACGCACCTCAAGCCCTA
 TGTCTTTCAGCAGCCCTGGCATGAGCAGTACCTGTTTCGACGTTTTTGTGTTTCTGCTGGGGC
 TCACTCTGGGGACTCTATGGCTTTGTGGGAAGCTGCTGGGCATGGCTGTCTGGTGGCTGCGT
 GGGGCCAGAAAGGTGAAGGAGACA**TAA**GGCCAGGTGCAGCCTTGGCGGGGTCTGTTTGGTGG
 GCGATGTCACCATTTCTAGGGAGCTTCCCACTAGTTCTGGCAGCCCCATTCTCTAGTCCTTC
 TAGTTATCTCCTGTTTTCTTGAAGAACAGGAAAAATGGCCAAAAATCATCCTTCCACTTGC
 TAATTTTGCTACAAATTCATCCTTACTAGCTCCTGCCTGCTAGCAGAAATCTTTCCAGTCCT
 CTTGTCTCCTTTGTTTGCCATCAGCAAGGGCTATGCTGTGATTCTGTCTCTGAGTGAAGTTG
 GACCACTGACCCTCAGATTTCCAGCCTTAAAATCCACCTTCCTTCTCATGCGCCTCTCCGAA
 TCACACCCTGACTCTTCCAGCCTCCATGTCCAGACCTAGTCAGCCTCTCTCACTCCTGCCCC
 TACTATCTATCATGGAATAACATCCAAGAAAGACACCTTGCATATTCTTTAGTTTCTGTTT
 TGTCTCCACATATTCTCTTCAATGCTCAGGAAGCCTGCCCTGTGCTTGAGAGTTTCAGGGC
 CGGACACAGGCTCACAGGTCTCCACATTGGGTCCCTGTCTCTGGTGCCACAGTGAGCTCCT
 TCTTGGCTGAGCAGGCATGGAGACTGTAGGTTTCCAGATTTCTGAAAAATAAAAGTTTACA
 GCGTTATCTCTCCCAACCTCACTAA

FIGURE 160

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71169

><subunit 1 of 1, 523 aa, 1 stop

><MW: 59581, pI: 8.68, NX(S/T): 1

MAGQRVLLLVGFLLPGVLLSEAAKILTISTVGGSHYLLMDRVSQILQDHGHNVTMLNHKRGPFMPDFKKEEKSYQVISWLAPEDHQREFKKSFDFFLEETLGGRGKFENLLNVLEYLALQCSHF
LNRKDIMDSLKNENFDMVIVETFDYCPFLIAEKLKPFVAILSTSFGSLEFGLPIPLSYVPV
FRSLTLDHMDFWGRVKNFLMFFSFCRRQQHMQSTFDNTIKEHFTEGSRPVLSHLLLKAELWF
INSDFAFDFARPLLNTVYVGGLMEKPIKVPVQDLENFIAKFGDSGFVLVTLGSMVNTCQNP
EIFKEMNNAFAHLPQGVIWKCQCShWPKDVHLAANVKIVDWLPQSDLLAHPSIRLFVTHGGQ
NSIMEAIQHGVPVMVGIPLEFGDQPENMVRVEAKKFGVSIQLKKLKAETLALKMKQIMEDKRYK
SAAVAASVILRSHPLSPTQRLVGWIDHVLQTGGATHLKPYPVFQQPWHEQYLFDFVFVLLGLT
LGTLLWLCGKLLGMAVWWLRGARKVKET

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 483-504

FIGURE 161

GGGCTGTTGATTGTGGGGGATTTTGAAGAGAGAGGAGGAATAGGAGGAAGGGGTTGAGGGGCT
GCCTCTGGCATATGCACACACTCACACATTCTGTACACCCCGTCACACACACATACCATGTT
CTCCATCCCCCAGGTCCAGCCCTCAGTGCTGTCCCATCCAGCAGGGGCTACCCTGAAGCTCT
GGCTGCAGCCCTCCCGTCCAGTGGGCAGGCGGCTTCATCCCTCCTTTCTCTCCCAAAGCCCCA
ACTGCTGTCACTGCATGCTCTGCCAAGGAGGAGGGAACTGCAGTGACAGCAGGAGTAAGAGT
GGGAGGCAGGACAGAGCTGGGACACAGGTATGGAGAGGGGGTTTCAGCGAGCCTAGAGAGGGC
AGACTATCAGGGGTGCCGGCGGTGAGAATCCAGGGAGAGGAGCGGAAACAGAAGAGGGGGCAGA
AGACCGGGGCACTTGTGGGTTGCAGAGCCCCCTCAGCCATGTTGGGAGCCAAGCCACACTGGC
TACCAGGTCCCCTACACAGTCCCGGGCTGCCCTTGGTTCTGGTGCTTCTGGCCCTGGGGGCC
GGGTGGGCCCAGGAGGGGTTCAGAGCCCGTCCTGCTGGAGGGGGAGTGCCTGGTGGTCTGTGA
GCCTGGCCGAGCTGCTGCAGGGGGGGCCCGGGGAGCAGCCCTGGGAGAGGCACCCCCTGGGC
GAGTGGCATTTGCTGCGGTCCGAAGCCACCACCATGAGCCAGCAGGGGAAACCGGCAATGGC
ACCAGTGGGGCCATCTACTTCGACCAGGTCTTGGTGAACGAGGGGCGGTGGCTTTGACCGGGC
CTCTGGCTCCTTCGTAGCCCCCTGTCCGGGGTGTCTACAGCTTCCGGTTCATGTGGTGAAGG
TGTACAACCGCCAAACTGTCCAGGTGAGCCTGATGCTGAACACGTGGCCTGTCATCTCAGCC
TTTGCCAATGATCCTGACGTGACCCGGGAGGCAGCCACCAGCTCTGTGCTACTGCCCTTGGA
CCCTGGGGACCGAGTGTCTCTGCGCCTGCGTCGGGGGAATCTACTGGGTGGTTGGAAATACT
CAAGTTTTCTCTGGCTTCCTCATCTTCCCTCTCTGAGAGGCCCAAGTCTTTCAAGCACAGAAT
CCAGCCCCCTGACAACTTTCTTCTGCCCTCTCTTGCCCCAGAAACAGCAGAGGCAGGAGAGAG
ACTCCCTCTGGCTCCTATCCCACCTCTTTGCATGGGACCCTGTGCCAAACACCCCAAGTTTAA
GAGAAGAGTAGAGCTGTGGCATCTCCAGACCAGGCCTTTCCACCCACCCACCCCAAGTTACC
CTCCCAGCCACCTGCTGCATCTGTTTCTGCTGCAGCCCTAGGATCAGGGCAAGGTTTGGA
AGAAGGAAGATCTGCACTACTTTGCGGCCTCTGCTCCTCCGGTTCACCCACCCCAAGTTCCT
GCTCAATGCTGATCAGGGACAGGTGGCGCAGGTGAGCCTGACAGGCCCCCACAGGAGCCAG
ATGGACAAGCCTCAGCGTACCCTGCAGGCTTCTTCTGTTGAGGAAAGCCAGCATCACGGATC
TCAGCCAGCACCGTCAGAAGCTGAGCCAGCACCGTATGGGCTAGGGTGGGAGGCTCAGCCAC
AGGCAGAAGGGTGGGAAGGGCCTGGAGTCTGTGGCTGGTGAGGAAGGAAGGAGGGTGTATTG
TCTAGACTGAACATGGTACACATTCTGCATGTATAGCAGAGCAGCCAGCAGGTAGCAATCCT
GGCTGTCCTTCTATGCTGGATCCCAGATGGACTCTGGCCCTTACCTCCCCACCTGAGATTAG
GGTGAGTGTGTTTGCTCTGGCTGAGAGCAGAGCTGAGAGCAGGTATACAGAGCTGGAAGTGG
ACCATGGAAAACATCGATAACCATGCATCCTCTTGCTTGGCCACCTCCTGAAACTGCTCCAC
CTTTGAAGTTTGAACTTTAGTCCCTCCACACTCTGACTGCTGCCTCCTTCCCTCCCAGCTCTC
TCACTGAGTTATCTTCACTGTACCTGTTCCAGCATATCCCCACTATCTCTCTTTCTCCTGAT
CTGTGCTGTCTTATTCTCCTCCTTAGGCTTCCTATTACCTGGGATTCCATGATTCATTCCTT
CAGACCCTCTCCTGCCAGTATGCTAAACCTCCCTCTCTCTTTCTTATCCCGCTGTCCCATT
GGCCCAGCCTGGATGAATCTATCAATAAAACAACCTAGAGAATGGTGGTCAGTGAGACACTAT
AGAATTACTAAGGAGAAGATGCCTCTGGAGTTTGGATCGGGTGTTACAGGTACAAGTAGGTA
TGTTGCAGAGGAAAATAAATATCAAACCTGTATACTAAAATTAAAAA

amino acids 1-32

FIGURE 163

[illegible]

FIGURE 164

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71184

><subunit 1 of 1, 388 aa, 1 stop

><MW: 43831, pI: 9.64, NX(S/T): 3

MKTLIAAYSGVLRGERQAEADRSQRSHGGPALSREGSGRWGTGSSILSALQDLFSVTWLNRS
KVEKQLQVISVLQWVLSFLVLGVACSAILMYIFCTDCWLIYFTWLVFDWNTPKKGGRRS
QWVRNWAVWRYFRDYFPIQLVKTHNLLTTRNYIFGYHPHGIMGLGAFCNFSTEATEVSKKFP
GIRPYLATLAGNFRMPVLREYLMSSGGICPVSRDTIDYLLSKNGSGNAIIIVVGAAESLSSM
PGKNAVTLRNRKGFVKLALRHGADLVPIYSFGENEVYKQVIFEEGSWGRWVQKKFQKYIGFA
PCIFHGRGLFSSDTWGLVPYSKPITTVVGEPITIPKLEHPTQQDIDLYHTMYMEALVKLFDK
HKTKFGLPETEVLEVN

Important features of the protein:**Transmembrane domain:**

amino acids 76-97

N-glycosylation sites.

amino acids 60-63, 173-176, 228-231

N-myristoylation sites.

amino acids 10-15, 41-46, 84-89, 120-125, 169-174, 229-234, 240-
245, 318-323, 378-383

FIGURE 165

GGGCGGCGGGATGGGGGCCGGGGCGGGCGGGCGCCGACTCGCTGAGGCCCCGACGCAGGGCCGGGCCGGGCCCA
 GGGCCGAGGAGCGCGGGCCAGAGCGGGGCCGGAGGCGACGCCGGGGACGCCCCGCGCGACGAGCAGGTGGCG
 GCGGCTGCAGGCTTGTCCAGCCGGAAGCCCTGAGGGCAGCTGTTCCCACTGGCTCTGCTGACCTTGTGCCTTGGA
 CGGCTGTCTCTCAGCGAGGGGCCGTGCACCCGCTCCTGAGCAGCGCC**CATG**GGCCTGCTGGCCTTCCTGAAGACCCA
 GTTCGTGCTGCACCTGCTGGTTCGGCTTTGTCTTCGTGGTGAGTGGTCTGGTCATCAACTTCGTCCAGCTGTGCAC
 GCTGGCGCTCTGGCCGGTCAGCAAGCAGCTCTACCGCCGCTCAACTGCCGCTCGCCTACTCACTCTGGAGCCA
 ACTGGTCATGCTGCTGGAGTGGTGGTCTGTCACGGAGTGACACTGTTACGGACCAGGCCACGGTAGAGCGCTT
 TGGGAAGGAGCACGCAGTCATCATCCTCAACCACAACCTTCGAGATCGACTTCCTCTGTGGGTGGACCATGTGTGA
 GCGCTTCGGAGTGCTGGGGAGCTCCAAGGTCTCGCTAAGAAGGAGCTGCTCTACGTGCCCCCTCATCGGCTGGAC
 GTGGTACTTTCTGGAGATTGTGTTCTGCAAGCGGAAGTGGGAGGAGGACCGGGACACCGTGGTGAAGGGCTGAG
 GCGCCTGTGCGACTACCCCGAGTACATGTGGTTTCTCCTGTACTGCGAGGGGACGCGCTTCACGGAGACCAAGCA
 CCGCGTTAGCATGGAGGTGGCGGCTGCTAAGGGGCTTCCTGTCTCAAGTACCACCTGCTGCCGCGGACCAAGGG
 CTTCAACCACCGCAGTCAAGTGCCTCCGGGGGACAGTCGCAGCTGTCTATGATGTAACCTGAACTTCAGAGGAAA
 CAAGAACCCGCTCCCTGCTGGGGATCCTCTACGGGAAGAAGTACGAGGCGGACATGTGCGTGAGGAGATTTCCTCT
 GGAAGACATCCCGCTGGATGAAAAGGAAGCAGCTCAGTGGCTTCATAAACTGTACCAGGAGAAGGACGCGCTCCA
 GGAGATATATAATCAGAAGGGCATGTTTCCAGGGGAGCAGTTTAAGCCTGCCCGGAGGCCGTGGACCCCTCCTGAA
 CTTCTGTCTCTGGGCCACCATTCTCCTGTCTCCCCTCTTCAGTTTTGTCTTGGGCGTCTTTGCCAGCGGATCACC
 TCTCCTGATCCTGACTTTCTTGGGGTTTGTGGGAGCAGCTTCCTTTGGAGTTCGCAGACTGATAGGAGAATCGCT
 TGAACCTGGGAGGTGGAGATTGCAG**TGAG**CTGAGATGGCATCACTGTACTCCAGCCTAGGCAACAGAGCAAGACT
 CAGTCTCAAAAAAAAAAAAAACAAAAAACCCAGAAATCTGGAGTTGAACTGTGTAGTTACTGACATGAAAA
 ATTCAGTAGAGGCTGAACAGCAGATTGAGCAGGCAGAAAAAATCAGCAAGCTTGAAGATGGTACCTTGAGATT
 TTTTCAAGCTAATGAAAAAGAATGAAGGAAAATTAACAGCCTCAGAGACCCATGGTGCACCGTCACACAAATCAA
 CATATGCATGATGAGAGTCCCAGAAGGAGAGGAGAGAAAGGGTCAGAAAGAATGGCCACAAGCTGATGAAAAACA
 GTAACCTACCCACTCAGGAAGCTCAGTGAACCTCCAATGAGGATGAATATCAGAGATCCACACCTAGATATTTTAT
 AATCAAAGTGTCAAATGACAAAGAATCTTGAAAGCAGCAAGAGATGAGCAACTTATCTTGTTCAAAGGATCTTTG
 ATCAGATTAACAGCTCATTTCTCCTCAGAAATCATGGGAGCCAGGAGATAGTGGGATGAACACTGTTGAAGGCAA
 AACCTTCAACTGTAATTATTGGACTTTTGGTCTTAGATGGTCTGACCTCTTTGTCTTCAGGGACAGTTTTTCA
 ATTTAATCCCTAATAACAATTAGTCAAGCTTCCTTGACCTGTAGGAAGGCCTGTCTTTAGGCCGGGCACAGTGGC
 TTACACCTGTAATCCCAGCACTTTGGGAGGCCAGACGGGTGGATCATTGGGGTCAGGCTGATCTCAAACCTCT
 GAGTTCAGGTGATCTGCCCGCCTCAGCCTCCCAAAGTGTTGTGATTGCAGGCGTGAGCCACTGCGCCTGGCCGGA
 ATTTCTTTTTAAGGCTGAATGATGGGGGCCAGGCACGATGGCTCACGCTGTGATCCCAAGTAGCTTGATTGTA
 AACATGCACCACCATGCCTGGCTAATTTTTGTATTTTTTAGTAGAGACGTGTTAGCCAGGCTGGTCTCGATCTCCT
 GACCTCAAGTGACCACCTGCCTCAGCCTCCCAAAGTACTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCTTGA
 GCATCTTGTGATGTGCTTATTGGCCATTTGTATATCTTCTATCTTCTTTGGGGAAATGTCTGTTCAAGTCCTTTG
 CCTTTTTTAAATTTTTATTATTTATTTATTTATTTTATTTTGGAGACAGGGTCTTGTCTGTTGCCAGGCTGGAGTA
 CAGTGGCACAGTCTTGGCTCACTGCAGCCTCGACCTCCTGGGCTGCAGTGATCCTCCACCTCAGCCTCCCTTGT
 AGCTGTATTTTTTTGTATTTTGTATTTTGTAGCTGTAGTTTTTGTATTTTTTGTGGAGACAGCATTTACCATGA
 TGCCCAGGCTGGTCTTGAACCTCTGAGCTCAAGTGATCTGCCTGCTTCAGCCTCCCAAAGTGCTGGGATTACAGA
 CATGAGCCACTGCACCTGGCAAACCTCCAAAATCAACACACACACAAAAAACACCTGATTCAAATGGGCA
 GAGGGGCCGGGTGTGGCCCCAACTACCAGGGAGACTGAAGTGGGAGGATCGCTTGGGCATGAGAAGTCGAGGCTG
 CAGTGAGTCGAGGTTGTGCGACTGCATTCCAGCCTGGACAACAGAGTGAGACCCTGTCTC

FIGURE 166

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71213

><subunit 1 of 1, 368 aa, 1 stop

><MW: 42550, pI: 9.11, NX(S/T): 1

MGLLAFLKTQFVLHLLVGFVFVVSGLVINFVQLCTLALWPVSKQLYRRRLNCRLAYSLWSQLV
MLLEWWSCTECTLFTDQATVERFGKEHAVIILNHNFEIDFLCGWTMCERFGVLGSSKVLAKK
ELLYVPLIGWTWYFLEIVFCKRKWEEDRDTVVEGLRRLSDYPEYMWFLLYCEGTRFTETKHR
VSMEVAAAKGLPVLKYHLLPRTKGFTTAVKCLRGTVAAVYDVTLNFRGNKNPSLLGILYGKK
YEADMCVRRFPLEDIPLDEKEAAQWLHKLYQEKDALQEIYNQKGMFPGEQFKPARRPWTLN
FLSWATILLSPLFSFVLGVFASGSPLLILTFLGFVGAASFGVRRLIGESLEPGRWRLQ

Important features of the protein:

Signal peptide:

amino acids 1-25

Transmembrane domains:

amino acids 307-323, 335-352

Tyrosine kinase phosphorylation sites.

amino acids 160-168, 161-169

Variable	Mean	SD	Min	Max
Age	31.15	4.12	18	45
Gender	0.48	0.50	0	1
Marital status	0.42	0.50	0	1
Education	12.15	1.12	10	15
Income	1.25	0.15	1.00	1.50
Health status	0.85	0.10	0.70	1.00
Work status	0.65	0.48	0	1
Family size	2.15	0.85	1	5
Religion	0.45	0.50	0	1
Political affiliation	0.35	0.48	0	1
Urban/rural	0.55	0.50	0	1
Employment type	0.45	0.50	0	1
Health insurance	0.75	0.43	0	1
Access to services	0.65	0.48	0	1
Health expenditure	0.15	0.05	0.00	0.30
Health status (self-rated)	0.85	0.10	0.70	1.00
Health status (physician-rated)	0.85	0.10	0.70	1.00
Health status (family-rated)	0.85	0.10	0.70	1.00
Health status (community-rated)	0.85	0.10	0.70	1.00
Health status (government-rated)	0.85	0.10	0.70	1.00
Health status (private-rated)	0.85	0.10	0.70	1.00
Health status (non-profit-rated)	0.85	0.10	0.70	1.00
Health status (for-profit-rated)	0.85	0.10	0.70	1.00
Health status (other-rated)	0.85	0.10	0.70	1.00
Health status (unrated)	0.85	0.10	0.70	1.00
Health status (not rated)	0.85	0.10	0.70	1.00
Health status (no rating)	0.85	0.10	0.70	1.00
Health status (no data)	0.85	0.10	0.70	1.00
Health status (missing)	0.85	0.10	0.70	1.00
Health status (unknown)	0.85	0.10	0.70	1.00
Health status (no info)	0.85	0.10	0.70	1.00
Health status (no eval)	0.85	0.10	0.70	1.00
Health status (no assess)	0.85	0.10	0.70	1.00
Health status (no check)	0.85	0.10	0.70	1.00
Health status (no test)	0.85	0.10	0.70	1.00
Health status (no exam)	0.85	0.10	0.70	1.00
Health status (no visit)	0.85	0.10	0.70	1.00
Health status (no consult)	0.85	0.10	0.70	1.00
Health status (no referral)	0.85	0.10	0.70	1.00
Health status (no advice)	0.85	0.10	0.70	1.00
Health status (no treatment)	0.85	0.10	0.70	1.00
Health status (no surgery)	0.85	0.10	0.70	1.00
Health status (no medication)	0.85	0.10	0.70	1.00
Health status (no therapy)	0.85	0.10	0.70	1.00
Health status (no diet)	0.85	0.10	0.70	1.00
Health status (no exercise)	0.85	0.10	0.70	1.00
Health status (no rest)	0.85	0.10	0.70	1.00
Health status (no sleep)	0.85	0.10	0.70	1.00
Health status (no stress)	0.85	0.10	0.70	1.00
Health status (no anxiety)	0.85	0.10	0.70	1.00
Health status (no depression)	0.85	0.10	0.70	1.00
Health status (no pain)	0.85	0.10	0.70	1.00
Health status (no fatigue)	0.85	0.10	0.70	1.00
Health status (no weakness)	0.85	0.10	0.70	1.00
Health status (no numbness)	0.85	0.10	0.70	1.00
Health status (no tingling)	0.85	0.10	0.70	1.00
Health status (no itching)	0.85	0.10	0.70	1.00
Health status (no burning)	0.85	0.10	0.70	1.00
Health status (no cold)	0.85	0.10	0.70	1.00
Health status (no flu)	0.85	0.10	0.70	1.00
Health status (no cough)	0.85	0.10	0.70	1.00
Health status (no sneeze)	0.85	0.10	0.70	1.00
Health status (no runny nose)	0.85	0.10	0.70	1.00
Health status (no sore throat)	0.85	0.10	0.70	1.00
Health status (no hoarse voice)	0.85	0.10	0.70	1.00
Health status (no dry mouth)	0.85	0.10	0.70	1.00
Health status (no dry eyes)	0.85	0.10	0.70	1.00
Health status (no dry skin)	0.85	0.10	0.70	1.00
Health status (no dry hair)	0.85	0.10	0.70	1.00
Health status (no dry nails)	0.85	0.10	0.70	1.00
Health status (no dry lips)	0.85	0.10	0.70	1.00
Health status (no dry feet)				

Variable	Mean	SD	Min	Max
Age	31.15	4.12	18	45
Gender	0.48	0.50	0	1
Marital status	0.42	0.50	0	1
Education	12.15	1.12	10	15
Income	1.25	0.15	1.00	1.50
Health status	0.85	0.10	0.70	1.00
Work status	0.65	0.48	0	1
Family size	2.15	0.85	1	5
Religion	0.45	0.50	0	1
Political affiliation	0.35	0.48	0	1
Urban/rural	0.55	0.50	0	1
Employment type	0.45	0.50	0	1
Health insurance	0.75	0.43	0	1
Access to services	0.65	0.48	0	1
Health expenditure	0.15	0.05	0.00	0.30
Health status (self-rated)	0.85	0.10	0.70	1.00
Health status (physician-rated)	0.85	0.10	0.70	1.00
Health status (family-rated)	0.85	0.10	0.70	1.00
Health status (community-rated)	0.85	0.10	0.70	1.00
Health status (government-rated)	0.85	0.10	0.70	1.00
Health status (private-rated)	0.85	0.10	0.70	1.00
Health status (non-profit-rated)	0.85	0.10	0.70	1.00
Health status (for-profit-rated)	0.85	0.10	0.70	1.00
Health status (other-rated)	0.85	0.10	0.70	1.00
Health status (unrated)	0.85	0.10	0.70	1.00
Health status (not rated)	0.85	0.10	0.70	1.00
Health status (no rating)	0.85	0.10	0.70	1.00
Health status (no data)	0.85	0.10	0.70	1.00
Health status (missing)	0.85	0.10	0.70	1.00
Health status (unknown)	0.85	0.10	0.70	1.00
Health status (no info)	0.85	0.10	0.70	1.00
Health status (no eval)	0.85	0.10	0.70	1.00
Health status (no assess)	0.85	0.10	0.70	1.00
Health status (no check)	0.85	0.10	0.70	1.00
Health status (no test)	0.85	0.10	0.70	1.00
Health status (no exam)	0.85	0.10	0.70	1.00
Health status (no visit)	0.85	0.10	0.70	1.00
Health status (no consult)	0.85	0.10	0.70	1.00
Health status (no referral)	0.85	0.10	0.70	1.00
Health status (no advice)	0.85	0.10	0.70	1.00
Health status (no treatment)	0.85	0.10	0.70	1.00
Health status (no surgery)	0.85	0.10	0.70	1.00
Health status (no medication)	0.85	0.10	0.70	1.00
Health status (no therapy)	0.85	0.10	0.70	1.00
Health status (no diet)	0.85	0.10	0.70	1.00
Health status (no exercise)	0.85	0.10	0.70	1.00
Health status (no rest)	0.85	0.10	0.70	1.00
Health status (no sleep)	0.85	0.10	0.70	1.00
Health status (no stress)	0.85	0.10	0.70	1.00
Health status (no anxiety)	0.85	0.10	0.70	1.00
Health status (no depression)	0.85	0.10	0.70	1.00
Health status (no pain)	0.85	0.10	0.70	1.00
Health status (no fatigue)	0.85	0.10	0.70	1.00
Health status (no weakness)	0.85	0.10	0.70	1.00
Health status (no numbness)	0.85	0.10	0.70	1.00
Health status (no tingling)	0.85	0.10	0.70	1.00
Health status (no itching)	0.85	0.10	0.70	1.00
Health status (no burning)	0.85	0.10	0.70	1.00
Health status (no cold)	0.85	0.10	0.70	1.00
Health status (no flu)	0.85	0.10	0.70	1.00
Health status (no cough)	0.85	0.10	0.70	1.00
Health status (no sneeze)	0.85	0.10	0.70	1.00
Health status (no runny nose)	0.85	0.10	0.70	1.00
Health status (no sore throat)	0.85	0.10	0.70	1.00
Health status (no hoarse voice)	0.85	0.10	0.70	1.00
Health status (no dry mouth)	0.85	0.10	0.70	1.00
Health status (no dry eyes)	0.85	0.10	0.70	1.00
Health status (no dry skin)	0.85	0.10	0.70	1.00
Health status (no dry hair)	0.85	0.10	0.70	1.00
Health status (no dry nails)	0.85	0.10	0.70	1.00
Health status (no dry lips)	0.85	0.10	0.70	1.00
Health status (no dry feet)				

FIGURE 168

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71234
><subunit 1 of 1, 143 aa, 1 stop
><MW: 15624, pI: 9.58, NX(S/T): 0
MHHS LQCPGAATRHIHL CVCFSFALALGHFLLISLVGKGLSLSCGVGGRQAGLR LIRPWVRR
EGKINFYTNGDSWGLRPASSVKFLGSAYTFFSLTWHTLLKASQGFSLFLGSKYLELQEPSWS
GPCPPGQLHCTCGVLLSFL
```

Important features of the protein:

Signal peptide:

amino acids 1-28

FIGURE 169

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGA
TTATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTTCAGCAACTAAAAAGCCAC
AGGAGTTGAACTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTAC
ATTAAAATCTGTTTTTTGTTCTCTTGTAAGTACGCTTTACCTTCCTAACACAGAGGATCTGT
CACTGTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCAC
ACCGTCCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTC
ACCAACTGTCTCACGTCTGGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGA
GCTGCGGCTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTT
CATAGGCGATGGCTCCCCTGCCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGG
CCAGGACGGGCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCCGCCAT
CTAACCTTTTCATGTCCTGCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCCAAGG
AAGCCAGAGCTTGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTT
TATGTGACAGGACTTGCAATTCCTGGAACATGAGGGGAACGCCGAGGAAAGCAAAGTGGA
GGGAAGGAACCTTGTGCCAAATTATGGGTGAGAAAAGATGGAGGTGTTGGGTTATCACAAGGC
ATCGAGTCTCCTGCATTCAGTGAGCATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACT
CGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGATCCACGTACCAGCTGCTG
AAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACCAG
CCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCTCCTTCCCTC
TGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGCTAATG
GCTCAGTGTTGGCCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCG
AACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCT
CAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGT
TCAAATGATCTCCAAGGGCCCTTATACCCCAGGAGACTTTGATTTGAATTTGAAACCCCAA
TCCAAACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATG
CCAACATTTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTG
GCCAACATGGTGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGT
GTGTGCCTGTATCCCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGT
GAAGGAGGCTGAGACAGGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGA
AAAAATAAAAAAAGAATTATGGTTATTTGTAA

172/249

FIGURE 170

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71277

><subunit 1 of 1, 109 aa, 1 stop

><MW: 11822, pI: 8.63, NX(S/T): 0

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTA
SPCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:

amino acids 1-15

172/249

FIGURE 171

GCGGGCCCGCGAGTCCGAGACCTGTCCCAGGAGCTCCAGCTCACGTGACCTGTCACTGCCTC
CCGCCGCCTCCTGCCCCGCGCC**ATG**ACCCAGCCGGTGCCCCGGCTCTCCGTGCCCCGCCGCGCT
GGCCCTGGGCTCAGCCGCACTGGGCGCCGCCTTCGCCACTGGCCTCTTCCTGGGGAGGCGGT
GCCCCCATGGCGAGGCCGCGAGAGCAGTGCCTGCTTCCCCCGAGGACAGCCGCCTGTGG
CAGTATCTTCTGAGCCGCTCCATGCGGGAGCACCCGGCGCTGCGAAGCCTGAGGCTGCTGAC
CCTGGAGCAGCCGCAGGGGGATTCTATGATGACCTGCGAGCAGGCCAGCTCTTGGCCAACC
TGGCGCGGCTCATCCAGGCCAAGAAGGCGCTGGACCTGGGCACCTTCACGGGCTACTCCGCC
CTGGCCCTGGCCCTGGCGCTGCCCCGCGGACGGGCGCGTGGTGACCTGCGAGGTGGACGCGCA
GCCCCCGGAGCTGGGACGGCCCCCTGTGGAGGCAGGCCGAGGCGGAGCACAAAGATCGACCTCC
GGCTGAAGCCCGCCTTGGAGACCCTGGACGAGCTGCTGGCGGCGGGCGAGGCCGGCACCTTC
GACGTGGCCGTGGTGGATGCGGACAAGGAGAACTGCTCCGCCTACTACGAGCGCTGCCTGCA
GCTGCTGCGACCCGGAGGCATCCTCGCCGTCCTCAGAGTCCTGTGGCGCGGGAAGGTGCTGC
AACCTCCGAAAGGGGACGTGGCGGCCGAGTGTGTGCGAAACCTAAACGAACGCATCCGGCGG
GACGTCAGGGTCTACATCAGCCTCCTGCCCCTGGGCGATGGACTCACCTTGGCCTTCAAGAT
CTAGGGCTGGCCCCTAGTGAGTGGGCTCGAGGGAGGGTTGCCTGGGAACCCAGGAATTGAC
CCTGAGTTTTAAATTCGAAAATAAAGTGGGGCTGGGACACAAAAAAAAAAAAAAAAAAAAA

FIGURE 172

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71282
><subunit 1 of 1, 262 aa, 1 stop
><MW: 28809, pI: 8.80, NX(S/T): 1
MTQPVPRLSVPAALALGSAALGAAFATGLFLGRRCPPWRGRREQCLLPPEDSRLWQYLLSRS
MREHPALRSLRLLTLEQPQGDMMTCEQAQLLANLARLIQAKKALDLGTFTGYSALALALAL
PADGRVVTCEVDAQPPPELGRPLWRQAEAEHKIDRLKPALETLDELLAAGEAGTFDVAVVDA
DKENCSAYYERCLQLLRPGGILAVLRVLWRGKVLQPPKGDVAAECVRNLNERIRRDVRVYIS
LLPLGDGLTLAFKI
```

Important features of the protein:**Signal peptide:**

amino acids 1-25

Transmembrane domains:

amino acids 8-30, 109-130

N-glycosylation site.

amino acids 190-193

Tyrosine kinase phosphorylation site.

amino acids 238-246

N-myristoylation sites.

amino acids 22-27, 28-33, 110-115, 205-210, 255-260

Amidation sites.

amino acids 31-34, 39-42

FIGURE 173

CCGCCGCCGCAGCCGCTACCGCCGCTGCAGCCGCTTTCCGCGGCCTGGGCCTCTCGCCGTCA
 GC**ATG**CCACACGCCTTCAAGCCCGGGGACTTGGTGTTTCGCTAAGATGAAGGGCTACCCTCAC
 TGGCCTGCCAGGATCGACGACATCGCGGATGGCGCCGTGAAGCCCCACCCAACAAGTACCC
 CATCTTTTTCTTTGGCACACACGAAACAGCCTTCCTGGGACCCAAGGACCTGTTCCCTACG
 ACAAATGTAAAGACAAGTACGGGAAGCCCAACAAGAGGAAAGGCTTCAATGAAGGGCTGTGG
 GAGATCCAGAACAACCCCCACGCCAGCTACAGCGCCCTCCGCCAGTGAGCTCCTCCGACAG
 CGAGGCCCCCGAGGCCAACCCCGCCGACGGCAGTGACGCTGACGAGGACGATGAGGACCGGG
 GGGTCATGGCCGTACAGCGGTAAACGCCACAGCTGCCAGCGACAGGATGGAGAGCGACTCA
 GACTCAGACAAGAGTAGCGACAACAGTGGCCTGAAGAGGAAGACGCCTGCGCTAAAGATGTC
 GGTCTCGAAACGAGCCCGAAAGGCCTCCAGCGACCTGGATCAGGCCAGCGTGTCCCCATCCG
 AAGAGGAGAACTCGGAAAGCTCATCTGAGTCGGAGAAGACCAGCGACCAGGACTTCACACCT
 GAGAAGAAAGCAGCGGTCCGGGCGCCACGGAGGGGGCCCTCTGGGGGGACGGAAAAAAGAA
 GCGCCGTCAGCCTCCGACTCCGACTCCAAGGCCGATTTCGGACGGGGCCAAGCCTGAGCCGG
 TGGCCATGGCGCGGTTCGGCGTCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCGACTCCGATGTG
 TCTGTGAAGAAGCCTCCGAGGGGGCAGGAAGCCAGCGGAGAAGCCTCTCCCGAAGCCGCGAGG
 GCGGAAACCGAAGCCTGAACGGCCTCCGTCCAGCTCCAGCAGTGACAGTGACAGCGACGAGG
 TGGACCGCATCAGTGAGTGGAAGCGGCGGGACGAGGCGCGGAGGCGCGAGCTGGAGGCCCGG
 CGGCGGCGAGAGCAGGAGGAGGAGCTGCGGCGCCTGCGGGAGCAGGAGAAGGAGGAGAAGGA
 GCGGAGGCGCGAGCGGGCCGACCGCGGGGAGGCTGAGCGGGGACGCGGCGGCAGCAGCGGGG
 ACGAGCTCAGGGAGGACGATGAGCCCGTCAAGAAGCGGGGACGCAAGGGCCGGGGCCGGGGT
 CCCCCGTCTCTCTGACTCCGAGCCCGAGGCCGAGCTGGAGAGAGAGGCCAAGAAATCAGC
 GAAGAAGCCGCAGTCTCAAGCACAGAGCCCGCCAGGAAACCTGGCCAGAAGGAGAAGAGAG
 TGCGGCCCCGAGGAGAAGCAACAAGCCAAGCCCGTGAAGGTGGAGCGGACCCGGAAGCGGTCC
 GAGGGCTTCTCGATGGACAGGAAGGTAGAGAAGAAGAAAGAGCCCTCCGTGGAGGAGAAGCT
 GCAGAAGCTGCACAGTGAGATCAAGTTTGCCCTAAAGGTGACAGCCCGGACGTGAAGAGGT
 GCCTGAATGCCCTAGAGGAGCTGGGAACCCTGCAGGTGACCTCTCAGATCCTCCAGAAGAAC
 ACAGACGTGGTGGCCACCTTGAAGAAGATTGCGCGTTACAAAGCGAACAAGGACGTAATGGA
 GAAGGCAGCAGAAGTCTATACCCGGCTCAAGTCGCGGGTCTCGGCCCAAAGATCGAGGCGG
 TGCAGAAAGTGAACAAGGCTGGGATGGAGAAGGAGAAGGCCGAGGAGAAGCTGGCCGGGGAG
 GAGCTGGCCGGGGAGGAGGCCCCCAGGAGAAGGCGGAGGACAAGCCCAGCACCGATCTCTC
 AGCCCCAGTGAATGGCGAGGCCACATCACAGAAGGGGGAGAGCGCAGAGGACAAGGAGCACG
 AGGAGGGTCGGGACTCGGAGGAGGGGCCAAGGTGTGGCTCCTCTGAAGACCTGCACGACAGC
 GTACGGGAGGGTCCCGACCTGGACAGGCCTGGGAGCGACCGGCAGGAGCGCGAGAGGGCACG
 GGGGACTCGGAGGCCCTGGACGAGGAGAGCT**TGA**GCCGCGGGCAGCCAGGCCCAGCCCCCGC
 CCGAGCTCAGGCTGCCCCTCTCCTTCCCCGGCTCGCAGGAGAGCAGAGCAGAGAACTGTGGG
 GAACGCTGTGCTGTTTGTATTTGTTCCCTTGGGTTTTTTTTCTCCTGCCTAATTTCTGTGATT
 TCCAACCAACATGAAATGACTATAAACGGTTTTTTAATGA

FIGURE 174

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71286

><subunit 1 of 1, 671 aa, 1 stop

><MW: 74317, pI: 7.61, NX(S/T): 0

MPHAFKPGDLVFAKMKGYPHWPARIDDIADGAVKPPPKNKYPIFFFFGTHETAFLGPKDLFPYD
KCKDKYGKPNKRKGFNEGLWEIQNNPHASYSAPPPVSSSDSEAPEANPADGSDADEDDEDRG
VMAVTAVTATAASDRMESDSDSDKSSDNSGLKRKTPALKMSVSKRARKASSDLDQASVSPSE
EENSESSSESEKTSQDFTPEKKA AVRAPRRGPLGGRKKKKAPSASDSDSKADSDGAKPEPV
AMARSASSSSSSSSSSSDSDSVSVKKPPRGRKPAEKPLPKPRGRKPKPERPPSSSSSDSDSDEV
DRISEWKRRDEARRRELEARRRREQEEELRRLREQEKEEKERRRERADRGEAERGSGGSSGD
ELREDDEPVKKRGRKGRGRGPPSSSDSEPEAELEAREAKKSAKKPQSSSTEPARKPGQKEKRV
RPEEKQQAKPVKVERTKRSEGFMDRKVEKKKEPSVEEKLQKLHSEIKFALKVDSPDVKRC
LNALEELGTLQVTSQILQKNTDVVATLKKIRRYKANKDVMKAAEVYTRLKSRVLGPKIEAV
QKVNKAGMEKEKAEEKLAGEELAGEEAPQEKAEDKPSTDLSAPVNGEATSQKGESAEDKEHE
EGRDSEEGPRCGSSEDLHDSVREGPDLD RPGSDRQERERARGDSEALDEES

Signal peptide:

amino acids 1-13

FIGURE 175

GTTGGTTCTCCTGGATCTTCACCTTACCAACTGCAGATCTTGGGACTCATCAGCCTCAATAATTATATTAAATTA
 ACACCATTTGAAAGAGAACATTGTTTTATC**ATG**AATGCTAATAAAGATGAAAGACTTAAAGCCAGAAGCCAAGA
 TTTTCACCTTTTTCTGCTTTGATGATGCTAAGCATGACCATGTTGTTTTCTTCCAGTCACTGGCACTTTGAAGCA
 AAATATTCCAAGACTCAAGCTAACCTACAAAGACTTGCTGCTTTCAAATAGCTGTATTCCCTTTTTGGGTTCATC
 AGAAGGACTGGATTTTCAAACCTCTTCTCTTAGATGAGGAAAGAGGCAGGCTGCTCTTGGGAGCCAAAGACCACAT
 CTTTCTACTCAGTCTGGTTGACTTAAACAAAAATTTTAAAGAAGATTTATTGGCCTGCTGCAAAGGAACGGGTGGA
 ATTATGTAAATTAGCTGGGAAAGATGCCAATACAGAATGTGCAAATTTTCATCAGAGTACTTCAGCCCTATAACAA
 AACTCACATATATGTGTGTGGAACCTGGAGCATTTCATCCAATATGTGGGTATATTGATCTTGGAGTCTACAAGGA
 GGATATTATATTCAAACCTAGACACACATAATTTGGAGTCTGGCAGACTGAAATGTCCTTTTCGATCCTCAGCAGCC
 TTTTGCTTCAGTAATGACAGATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAAGATACTGCATT
 CACTCGATCCCTTGGGCCTACTCATGACCACCACTACATCAGAACTGACATTTTCAGAGCACTACTGGCTCAATGG
 AGCAAAATTTATTGGAACCTTTCTTCATACCAGACACCTACAATCCAGATGATGATAAAATATATTTCTTCTTTTCG
 TGAATCATCTCAAGAAGGCAGTACCTCCGATAAAACCATCCTTTCTCGAGTTGGAAGAGTTTGTAAAGAATGATGT
 AGGAGGACAACGCAGCCTGATAAACAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTCCTGGAAG
 TGATGGGGCAGATACTTACTTTGATGAGCTTCAAGATATTTATTTACTCCCCACAAGAGATGAAAGAAATCCTGT
 AGTATATGGAGTCTTTACTACAACCAGCTCCATCTTCAAAGGCTCTGCTGTTTGTGTGTATAGCATGGCTGACAT
 CAGAGCAGTTTTTAATGGTCCATATGCTCATAAGGAAAGTGCAGACCATCGTTGGGTGCAGTATGATGGGAGAAT
 TCCTTATCCACGGCCTGGTACATGTCCAAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGA
 TGATGTCATCAGTTTCATAAAGCGGCACCTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGAGGACCAACGTT
 CAAGAGAATCAATGTGGATTACAGACTGACACAGATAGTGGTGGATCATGTTCATTGCAGAAGATGGCCAGTACGA
 TGTAATGTTTTCTTGGAAACAGACTTGGAACTGTCTCAAAGTTGTGAGCATTTCAAAGGAAAAGTGGAAATATGGA
 AGAGGTAGTGTGGAGGAGTTGCAGATATTCAAGCACTCATCAATCATCTTGAACATGGAATTGTCTCTGAAGCA
 GCAACAATTGTACATTGGTTCCCGAGATGGATTAGTTGAGCTCTCCTTGCACAGATGCGACACTTATGGGAAAGC
 TTGCGCAGACTGTTGTCTTGCCAGAGACCCCTACTGTGCCTGGGATGGAATGCATGCTCTCGATATGCTCCTAC
 TTCTAAAAGGAGAGCTAGACGCCAAGATGTAAATATGGCGACCCAATCACCCAGTGTGGGACATCGAAGACAG
 CATTAGTCATGAAACTGCTGATGAAAAGGTGATTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATACC
 TAAATCCCAACAAGCAACTATTAAATGGTATATCCAGAGGTGAGGGATGAGCATCGAGAGGAGTTGAAGCCCCGA
 TGAAGAATCATCAAAACGGAATATGGGCTACTGATTGCAAGTTTGCAGAAGAAGGATTCTGGGATGTATTACTG
 CAAAGCCCAGGAGCACACTTTTCATCCACACCATAGTGAAGCTGACTTTGAATGTCATTGAGAATGAACAGATGGA
 AAATACCCAGAGGGCAGAGCATGAGGAGGGGCAGGTCAAGGATCTATTGGCTGAGTCACGGTTGAGATACAAAGA
 CTACATCAAATCCTTAGCAGCCCCAACTTCAGCCTCGACCAGTACTGCGAACAGATGTGGCACAGGGGAGAAGCG
 GAGACAGAGAAAACAAGGGGGGCCAAAGTGGGAAGCACATGCGAGAAATGAAGAAGAAACGAAATCGAAGACATCA
 CAGAGACCTGGATGAGCTCCCTAGAGCTGTAGCCACG**TAG**TTTTCTACTTAATTTAAAGAAAAGAAATTCCTTACC
 TATAAAAACATTGCCTTCTGTTTTGTATATCCCTTATAGTAATTCATAAATGCTTCCCATGGAGTTTTGCTAAGG
 CACAAGACAATAATCTGAATAAGACAATATGTGATGAATATAAGAAAGGGCAAAAATTCATTTGAACCAGTTTT
 CCAAGAACAATCTTGCACAAGCAAAAGTATAAGAATTATCCTAAAAATAGGGGGTTTACAGTTGTAAATGTTTTA
 TGTTTTGAGTTTTTGAATTTATTGTTCATGTAAATAGTTGAGCTAAGCAAGCCCCGAATTTGATAGTGTATAAGGT
 GCTTTATTCCCTCGAATGTCCATTAAAGCATGGAATTTACCATGCAGTTGTGCTATGTTCTTATGAACAGATATAT
 CATTCTATTGAGAACCAGCTACCTTGTGGTAGGGAATAAGAGGTGAGACACAAATTAAGACAACCTCCCATATC
 AACAGGAACCTTTCTCAGTGAGCCATTCACTCCTGGAGAATGGTATAGGAATTTGGAGAGGTGCATTATTTCTTTC
 TGGCCACTGGGGTTAAATTTAGTGTACTACAACATTGATTTACTGAAGGGCACTAATGTTTCCCCCAGGATTTCT
 ATTGACTAGTCAGGAGTAACAGGTTACAGAGAGAAGTTGGTGCTTAGTTATGTGTTTTTTAGAGTATATACTAA
 GCTCTACAGGGACAGAATGCTTAATAAATACTTTAATAAGATATGGGAAAATATTTTAATAAAACAAGGAAAAACA
 TAATGATGTATAATGCATCCTGATGGGAAGGCATGCAGATGGGATTTGTTAGAAGACAGAAGGAAAGACAGCCAT
 AAATTCTGGCTTTGGGGAAGAACTCATATCCCCATGAAAAGGAAGAACAATCACAAATAAAGTGAGAGTAATGTAA
 TGGAGCTCTTTTCACTAGGGTATAAGTAGCTGCCAATTTGTAATTCATCTGTTAAAAAAATCTAGATTATAACA
 AACTGCTAGCAAAATCTGAGGAACATAAATTTCTCTGAAGAATCATAGGAAGAGTAGACATTTTATTATAACC
 AATGATATTTTCAGTATATATTTTCTCTCTTTTAAAAAATATTTATCATACTCTGTATATTATTTCTTTTTACTGC
 CTTTATTCTCTCCTGTATATTGGATTTTGTGATTATATTTGAGTGAATAGGAGAAAACAATATATAACACACAGA
 GAATTAAGAAAATGACATTTCTGGGGAGTGGGGATATATATTTGTTGAATAACAGAACGAGTGTAATAATTTTAAC
 AACGGAAGGGTTAAATTAACCTTTTGACATCTTCACTCAACCTTTTCTCATTGCTGAGTTAATCTGTTGTAATT
 GTAGTATTGTTTTTGAATTTAACAATAAATAAGCCTGCTACATGT

FIGURE 176

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71883

><subunit 1 of 1, 777 aa, 1 stop

><MW: 89651, pI: 7.97, NX(S/T): 3

MNANKDERLKARSQDFHLFPALMMLSMTMLFLPVGTGLKQNIPLRLKLTYSKDLLLSNSCIPFL
 GSSEGLDFQTLTLLDEERGRLLLGAKDHIFLLSLVDLNKNFKKIYWPAAKERVELCKLAGKDA
 NTECANFIRVLQPYNKTHIYVCGTGAFHPICGYIDLGVYKEDIIFKLDTHNLESGRLKCPFD
 PQQPFASVMTDEYLYSGTASDFLGKDTAFTRSLGPTHDDHYIRTDISEHYWLNKAKFIGTFF
 IPDTYNPDDDKIYFFFRESSQEGSTSDKTILSRVGRVCKNDVGGQSLINKWTTFLKARLIC
 SIPGSDGADTYFDELQDIYLLPTRDERNPVVYGVFTTTSSIFKGSAVCVYSMADIRAVFNGP
 YAHKESADHRWVQYDGRIPIPRPGTCPSKTYDPLIKSTRDFPDDVISFIKRHSVMYKSVYPV
 AGGPTEFKRINVDYRLTQIVVDHVIAEDGQYDVMFLGTDIGTVLKVVVISKEKWNMEEVVLEE
 LQIFKHSSIIILNMELSLKQQQLYIGSRDGLVQLSLHRCDTYKGACADCCLARDPYCAWDGNA
 CSRYAPTSKRRARRQDVKYGPITQCWDIEDSISHETADEKVIFGIEFNSTFLECIPKSQQA
 TIKWYIQRSGDEHREELKPDERIIKTEYGLLIRSLQKKDSGMYECKAQEHTFIHTIVKLTLN
 VIENEQMENTQRAEHEEGQVKDLLAESRLRYKDYIQILSSPNFSLDQYCEQMWHREKRRQRN
 KGGPKWKHMQEMKKRNRHRDLDELPRAVAT

Important features of the protein:

Signal peptide:

amino acids 1-36

N-glycosylation sites.

amino acids 139-142, 607-610, 724-727

Tyrosine kinase phosphorylation site.

amino acids 571-576

Gram-positive cocci surface proteins 'anchoring' hexapeptide.

amino acids 32-37

FIGURE 177

CCCTGACCTCCCTGAGCCACACTGAGCTGGAAGCCGAGAGGTTCATCCTGGAGCATGCCACCCGCGGGGAGCAGA
 CAACCTCCCAGGTAAGCTGGGAGCAAGACCTGAAGCTGTTTCTTCAGGAGCCTGGTGTATTTTCCCCACCCAC
 CTCAGCAGTTTCAGCCAGCAGGGACTGATCAGGTGTGTCTCCTGGAGTGGGAGCAGAAGGCGTGGCTGGCAAGA
 GTGGCCTGGAGAAAGAGGTTTCAGCGCTTGACCAGCCGAGCTGCCCCGTGACTACAAGATCCAGAACCATGGGCATC
 GGGTGAGGTGGGGGGGCACAGGTGTCTGTGCACCTTCTTGTCTCAGCAAGAAGAGCTGAGAGAGGGGATCTTGG
 AGCCATTGAGGGTGTCTGAGGCTACAGAGGGGAGGGAAAGGTATTTAAGGTAACAGTGTGGCACAATAGTTAA
 GAGCACAGTTTTTGGAGCTAGACCGACATAGGTTCAAATTCTCTTCTGTTGCTTCTAGTTCTGTAGCCCCAGGT
 AAGGGAGTGACTTAACCTCTCTGGACTTCAATTTCTCATCACTAAAGTAGGGCCAATAATAGCACCCACCTCAT
 AGGGAAGATTAAATGACATAATGTATGTG**ATG**CAACTAGCAAAGTACCAGTCCCATAGTAAGTCATGCCCCACAG
 TATTTCCACCCACCCCTGTTCTCTGCCTTCCCAACCAGGTAAGTGAACGACTGGAGCAGAGGCGGCAGCAGGCTT
 CAGAGCGGGAGGCTCCAAGCATAGAACAGAGGTTACAGGAAGTGCGAGAGAGCATCCGCCGGGCACAGGTGAGCC
 AGGTGAAGGGGGCTGCCCGCTGGCCCTGCTGCAGGGGGCTGGCTTAGATGTGGAGCGCTGGCTGAAGCCAGCCA
 TGACCCAGGCCCAGGATGAGGTGGAGCAGGAGCGGCGGCTCAGTGAGGCTCGGCTGTCCAGAGGGACCTCTCTC
 CAACCGCTGAGGATGCTGAGCTTTCTGACTTTGAGGAATGTGAGGAGACGGGAGAGCTCTTTGAGGAGCCTGCC
 CCAAGCCCTGGCCACGAGGGCCCTCCCTGCCCTGCACACGTGGTATTTTCGCTATCAGGCAGGGCGTGAGGATG
 AGCTGACAATCACGGAGGGTGAGTGGCTGGAGGTATAGAGGAGGGAGATGCTGACGAATGGGTCAAGGCTCGGA
 ACCAGCACGGCGAGGTAGGCTTTGTCCCTGAGCGATATCTCAACTTCCCGGACCTCTCCCTCCCAGAGAGCAGCC
 AAGACAGTGACAATCCCTGCGGGGCAGAGCCACAGCATTCTGGCACAGGCCCTGTACAGCTACACCGGACAGA
 GTGCAGAGGAGCTGAGCTTCCCTGAGGGGGCACTCATCCGTCTGCTGCCCCGGGCCAAGATGGAGTAGATGACG
 GCTTCTGGAGGGGAGAATTTGGGGGCCGTGTTGGGGTCTTCCCTCCCTGCTGGTGGAAGAGCTGCTTGGCCCC
 CAGGGCCACCTGAACTCTCTGACCCTGAACAGATGCTGCCGTCCCTTCTCCTCCAGCTTCTCCCCACCTGCAC
 CTACCTCTGTGTTGGATGGGGCCCTGCACCTGTCTGCTGGGACAAAGCCCTGGACTTCCCTGGGTCTCTGG
 ACATGATGGCACCTCGACTCAGGCCGATGCGTCCACCACCTCCCCCGCCGGCTAAAGCCCCGGATCCTGGCCACC
 CAGATCCCCCTCACCT**TGA**AGGCCAGGGAAGCCTTGACCCCCAGTGATGCTGCTGTCCCTATCTTCAAGCTGTGAGA
 CCACACCATCAATGATCCAGAGCAACACAGCCAAAAGCTGGAATCGCCCTTATTTCCACCCTCACCTCCAAGGGT
 GGAAACTTGCCCCCTCCCATTTCTAGAGCTGGAACCCACTCCTTTTTTCCATTGTTCTATCATCTCTAGGACC
 GGAATACTACCTTCTCTTCTGTCTGATGACCCTATCTAGGGTGAGTGAATGCCTGAAATCTCTGGGGCTGGAACC
 ATCCATCAAGGTCTCTAGTAGTCTTGGCCACCTCTTCCCCACCCTGGCTCCATGACCCACCCCACTCTGGATG
 CCAGGGTCACTGGGGTTGGGCTGGGAGAGGAACAGGCCCTTGGGAATCAGGAGCTGGAGCCAGGATGCGAAGCAG
 CTGTAATGGTCTGAGCGGATTTATTGACAATGAATAAAGGGCACGAAGGCCAGGCCAGGGCTGGGCTCTTGTG
 CTAAGAGGGCAGGGGGCCTACGGTGCTATTGCTTTAGGGGCCACCACGGGCAGGGGCCCTGCTCCCAGCTGCCAC
 GCTCTATCATATGGAGCGAGGTGTTGGGGAAGCGGGCAGGCAGCCTGTTGCAGGCAGGGGAAGGAGAAGAGAC
 TGAGGGGCTGTGACCTCTCCTGAGGCCCCAGCCTGAGACTGTGCAACTCCAGGTGGAAGTAGAGCTGGTCCCTC
 AGCTGGGGGGCAGTGCTGTCCAGTGGAGGGGAGGGCTTTCACGCCCCACCCACCCCTGGCCCTGCCAGCTGGTAG
 TCCATCAGCACAAATGAAGGAGACTTGGAGAAGAGGAAGAATAACACTGTTGCTTCTTCAAGCTGTGTCCAGC
 TTTTCCCCTGGGGCTCCAGGACCTTCCCTACCTCCACCACCAACCAAGGGATTTATAGCAAAGGCTAAGCCTGC
 AGTTTACTCTGGGGGTTCCAGGGAGCCGAAAGGCTTAAATAGTTTAAAGTAGGTGATGGGAAGATGAGATTACCTCA
 TTTAGGGCTCAGGCAGACTCACCTCACATACTCCCTGCTCCCTGTGGTAGAGACACCTGAGAGAAAGGGGAGGGG
 TCAACAATGAGAGACCAGGAGTAGGTCTATCAGTGCCCCCAGAGTAGAGAGCAATAAGAGCCCAGCCCAGTGC
 AGTCCCGGCTGTGTTTTCTACCTGGTGATCAGAAGTGTCTGGTTTGGCTGGCTGCCATTTGCCTCTTGAGTGG
 GCAGCCCTGGGCTTGGGCCCCCTCCCTCCGGCCCTCAGTGTTGGCTCTGCAGAAGCTCTGGGGTTCCTTCAAGTG
 CACGAGGGGTAGGCTGCTGTCCCTGAGTCTCCATCTGTACTGGGGGGCTGGCTAGGACCTGGGGCTGTGGCC
 TCTCAGGGGGCAGCCTCTCCATGGCAGGCATCCCTGCCTTGGGCTGCCCTCCCCAGACCCCTGACCACCCCTG
 GGTCTGTCCCCCACCAGAGCCCCAGCTCCTGTCTGTGGGGAGCCATCACGGTGTTCTGTGAGTCCATAGCGCT
 TCTCAATGTGTGTACCCGGAACCTGGGAGGGGAGGGAACACTGGGGTTTAGGACCACAACCTCAGAGGCTGCTTG
 GCCCTCCCCTCTGACCAGGGACATCCTGAGTTTGGTGGCTACTTCCCTCTGGCCTAAGGTAGGGGAGGCCTTCTC
 AGATTGTGGGGCACATTGTGTAGCCTGACTTCTGCTGGAGCTCCAGTCCAGGAGGAAAGAGCCAAGGCCCACTT
 TTGGGATCAGGTGCCTGATCACTGGGCCCCCTACCTCAGCCCCCTTTCCCTGGAGCACCTGCCCCACCTGCCCA
 CAGAGAACACAGTGGTCTCCCTGTCCGGGGCGGCTTTTCTTCTTGGAGCGTCCCTGACGGACAAGTGGAG
 GCCTCTTGCTGCGGCTGCAATGGATGCAAGGGGCTGCAGAGCCCAGGTGCACTGTGTGATGATGGGAGGGGCTC
 CGTCTGCAGGCTGGAGGTGGCATCCACACTGGACAGCAGGAGGGAGGTGAGGGTAACATTTCCATTTCCCT
 TCATGTTTTGTTTCTTACGTTCTTTCAGCATGCTCCTTAAACCCCAAGCCCCAATTTCCCCAAGCCCCATTT
 TTTCTTGTCTTTATCTAATAAACTCAATATTAAAG

FIGURE 178

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73401
```

><subunit 1 of 1, 370 aa, 1 stop

><MW: 40685, pI: 4.53, NX(S/T): 0

MQLAKYQSHSKSCPTVFPPTPVLCLPNQVLQRLEQRRQQASEREAPSIEQRLQEVRESIRRA
 QVSQVKGAARLALLQGAGLDVERWLKPAMTQAQDEVEQERRLSEARLSQRDLSPTAEDAELS
 DFEECEETGELFEEPAPQALATRALPCPAHVVFQAGREDELTITEGEWLEVIEEGDADEW
 VKARNQHGEVGVFPERYLNFPDLSLPSSQSDSNPCGAEPATAFLAQALYSYTGQSAEELSFP
 EGALIRLLPRAQDGVDDGFWRGEFGGRVGVFPSLLVEELLGPPGPPPELSDPEQMLPSPSPPS
 FSPPAPTSLVDGPPAPVLPDGDKALDFPGFLDMMAPRLRPMRPPPPPPAKAPDPGHPDPLT

FIGURE 179A

CACAGGGAGACCCACAGACACATATGCACGAGAGAGACAGAGGAGGAAAGAGACAGAGACAAAGGCACAGCGGAA
GAAGGCAGAGACAGGGCAGGCACAGAAGCGGCCAGACAGAGTCTACAGAGGGAGAGGCCAGAGAAGCTGCAGA
AGACACAGGCAGGGAGAGACAAAGATCCAGGAAAGGAGGGCTCAGGAGGAGAGTTTGGAGAAGCCAGACCCCTGG
GCACCTCTCCCAAGCCCAAGGACTAAGTTTTCTCCATTTCTTTAACGGTCTCAGCCCTTCTGAAAACCTTTGCC
TCTGACCTTGGCAGGAGTCCAAGCCCCAGGCTACAGAGAGGAGCTTTCCAAAGCTAGGGTGTGGAGGACTTGGT
GCCCTAGACGGCCTCAGTCCCTCCCAGCTGCAGTACCAGTGCC**ATGT**CCCAGACAGGCTCGCATCCCGGGAGGGG
CTTGGCAGGGCGCTGGCTGTGGGGAGCCCAACCTGCCTCCTGCTCCCCATTGTGCCGCTCTCCTGGCTGGTGTG
GCTGCTTCTGCTACTGCTGGCCTCTCTCCTGCCCTCAGCCCGCTGGCCAGCCCTCCCCGGGAGGAGGAGAT
CGTGTTCAGAGAAGCTCAACGGCAGCGTCTGCTGGCTCGGGCGCCCTGCCAGGCTGTTGTGCCGCTTGCA
GGCCTTTGGGGAGACGCTGCTACTAGAGCTGGAGCAGGACTCCGGTGTGCAGGTGAGGGGGCTGACAGTGCAGTA
CCTGGGGCAGGCGCCTGAGCTGCTGGGTGGAGCAGAGCCTGGCACCTACCTGACTGGCACCATCAATGGAGATCC
GGAGTCGGTGGCATCTCTGCACTGGGATGGGGGAGCCCTGTTAGGCGTGTACAAATATCGGGGGGCTGAACTCCA
CCTCCAGCCCTGGAGGGAGGCACCCCTAACTCTGCTGGGGGACCTGGGGCTCACATCCTACGCCGGAAGAGTCC
TGCCAGCGGTCAAGGTCCCATGTGCAACGTCAAGGCTCCTCTTGAAGCCCCAGCCCCAGACCCGAAGAGCCAA
GCGCTTTGCTTCACTGAGTAGATTTGTGGAGACACTGGTGGTGGCAGATGACAAGATGGCCGCATTCCACGGTGC
GGGGCTAAAGCGCTACCTGCTAACAGTGATGGCAGCAGCAGCCAAGGCCTTCAAGCACCCAAGCATCCGCAATCC
TGTCAGCTTGGTGGTGAAGTCCGCTAGTGATCCTGGGGTCAAGCGAGGAGGGGCCCCAAGTGGGGCCCAGTGCTGC
CCAGACCCTGCGCAGCTTCTGTGCCTGGCAGCGGGGCCTCAACACCCCTGAGGACTCGGGCCCTGACCACCTTTGA
CACAGCCATTCTGTTTACCCGTGAGGACCTGTGTGGAGTCTCCACTTGCACACGCTGGGTATGGCTGATGTGGG
CACCGTCTGTGACCCCGCTCGGAGCTGTGCCATTGTGGAGGATGATGGGCTCCAGTCAGCCTTCACTGCTGCTCA
TGAAGTGGGTGATGTCTTCAACATGCTCCATGACAACCTCAAGCCATGCATCAGTTTGAATGGGCCTTTGAGCAC
CTCTCGCCATGTGATGGCCCTGTGATGGCTCATGTGGATCCTGAGGAGCCCTGGTCCCCCTGCAGTGGCCCTT
CATCACTGACTTCTTGACAATGGCTATGGGCACTGTCTCTTAGACAAACCAGAGGCTCCATTGCATCTGCCTGT
GACTTTCCCTGGCAAGGACTATGATGCTGACCGCCAGTGCCAGCTGACCTTCGGGCCCGACTCACGCCATTGTCC
ACAGCTGCCGCCGCCCTGTGCTGCCCTCTGGTGTCTGGCCACCTCAATGGCCATGCCATGTGCCAGACCAAAACA
CTCGCCCTGGGCCGATGGCACACCCCTGCGGGCCCGCACAGGCCATGCATGGGTGGTGGTGGCTGCCATGGACCA
GCTCCAGGACTTCAATATTCCACAGGCTGGTGGCTGGGGTCCCTGGGGACCATGGGGTGAAGTGTCTCGGACCTG
TGGGGGTGGTGTCCAGTTCTCCTCCCGAGACTGCACGAGGCTGTCCCCCGGAATGGTGGCAAGTACTGTGAGGG
CCGCCGTACCCGCTTCCGCTCCTGCAACACTGAGGACTGCCCAACTGGCTCAGCCCTGACCTTCCGCGAGGAGCA
GTGTGCTGCCTACAACCACCGCACCGACCTCTTCAAGAGCTTCCCAGGGCCCATGGACTGGGTTCCTCGCTACAC
AGGCGTGGCCCCCAGGACCAAGTGAACACTCACCTGCCAGGCCCGGGCACTGGGCTACTACTATGTGCTGGAGCC
ACGGGTGGTAGATGGGACCCCTGTTCCCGGACAGCTCCTCGGTCTGTGTCAGGGCCGATGCATCCATGCTGG
CTGTGATCGCATCATTGGCTCCAAGAAGAAGTTTGACAAGTGATGGTGTGCGGAGGGGACGGTTCTGGTTGCAG
CAAGCAGTCAGGCTCCTTCAGGAAATTCAGGTACGGATAACAATGTGGTCACTATCCCCGCGGGGGGCCACCCA
CATTCTTGTCCGGCAGCAGGGAAACCTGGCCACCGGAGCATCTACTTGGCCCTGAAGCTGCCAGATGGCTCCTA
TGCCCTCAATGGTGAATACAGCTGATGCCCTCCCCACAGATGTGGTACTGCCTGGGGCAGTCAGCTTGCCTA
CAGCGGGCCCACTGCAGCCTCAGAGACACTGTGAGCCATGGGCCACTGGCCAGCCTTTGACACTGCAAGTCTT
AGTGGCTGGCAACCCCAAGGACACAGCCTCCGATACAGCTTCTCGTGCCCCGCGCCGACCCCTTCAACGCCACG
CCCCACTCCCCAGGACTGGCTGCACCGAAGAGCACAGATTCTGGAGATCCTTCGGCGGGCGCCCTGGCGGGCAG
GAA**ATAA**CTCACTATCCCGGCTGCCCTTTCTGGGCACCGGGGCTCGGACTTAGCTGGGAGAAAGAGAGAGCTT
CTGTTGCTGCCTCATGCTAAGACTCAGTGGGGAGGGGCTGTGGGCGTGAGACCTGCCCTCCTCTCTGCCCTAAT
GCGCAGGCTGGCCCTGCCCTGGTTTCTGCCCCTGGGAGGAGTATGGGTAGTGGATGGAAGGGGCTGACAGAC
AGCCCTCATCTAAACTGCCCTCTGCCCTGCGGGTCAAGGAGGGAGGGGGAAGGCAGGGAGGGGCTGGGGCC
CAGTTGTATTTATTTAGTATTTATTTCACTTTTATTTAGCACCAAGGGAAGGGGACAAGGACTAGGGTCTGGGGAA
CCTGACCCCTGACCCCTCATAGCCCTCACCCCTGGGGCTAGGAAATCCAGGGTGGTGGTGATAGGTATAAGTGGTG
TGTGTATGCGTGTGTGTGTGTGTGTGAAATGTGTGTGTGCTTATGTATGAGGTACAACCTGTTCTGCTTTCTC
TTCTGAATTTTATTTTGGGAAAAGAAAAGTCAAGGTAGGGTGGGCCCTTCAGGGAGTGAGGGATTATCTTTT
TTTTTTTTTCTTTCTTTCTTTCTTTTGTGAGACAGAATCTCGCTCTGTGCGCCAGGCTGGAGTGCAATG
GCACAATCTCGGCTCACTGCATCCTCCGCTCCCGGGTCAAGTGATTCTCATGCCTCAGCCTCCTGAGTAGCTG
GGATTACAGGCTCCTGCCACCACGCCAGCTAATTTTGTGTTTGTGTTTGGAGACAGAGTCTCGTATTGTC
ACCAGGGCTGGAATGATTTTCACTCACTGCAACCTTCGCCACCTGGGTTCAGCAATTCTCCTGCCTCAGCCTCC
CGAGTAGCTGAGATTATAGGCACCTACCACCACGCCCGCTAATTTTGTATTTTGTAGAGACGGGGTTTCAC
CATGTTGGCCAGGCTGGTCTCGAAGCTCCTGACCTTAGGTGATCCACTCGCCTTCATCTCCAAAGTGCTGGGATT
ACAGGCGTGAGCCACCGTGCTGGCCACGCCCACTAATTTTGTATTTTGTAGAGACAGGGTTTACCATGT
TGGCCAGGCTGCTTGAAGCTCCTGACCTCAGGTAATCGACCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG
TGTGAGCCACCACGCCCGGTACATATTTTAAATTGAATTCTACTATTTATGTGATCCTTTTGGAGTCAGACAG

[illegible]

FIGURE 180

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73492

><subunit 1 of 1, 837 aa, 1 stop

><MW: 90167, pI: 8.39, NX(S/T): 1

MSQTGSHPGRGLAGRWLWGAQPCLLLPIVPLSWLVWLLLLLLASLLPSARLASPLPREEEIV
 FPEKLNGSVLPGSGAPARLLCRLQAFGETLLLELEQDSGVQVEGLTVQYLGQAPELLGGAEP
 GTYLTGTINGDPESVASLHWDGGALLGVLYRGAEHLQPLEGGTPNSAGGPGAHILRRKSP
 ASGQGPMCNCVAPLGSPSPRRRAKRFAFSLSRFVETLVVADDKMAAFHGAGLKRYLLTVMAA
 AAKAFKHPSIRNPVSLVVTRLVILGSGEEGPQVGPSAAQTLRSFCAWQRLNTPEDSGPDHF
 DTAILFTRQDLGCVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAHELGHVFNMLHD
 NSKPCISLNGPLSTSRHVMAPVMAHVDPEEPWSPCSARFITDFLDNGYGHCLLDKPEAPLHL
 PVTFFPGKDYDADRQCQLTFGPDSRHCPQLPPPCAALWCSGHLNGHAMCQTKHSPWADGTPCG
 PAQACMGGRCLHMDQLQDFNIPQAGGWGPWGPWGDCSRTC GGGVQFSSRDCTRVPVRNGGKY
 CEGRRTRFRSCNTEDCPTGSALTFRREEQCAAYNHRTDLFKSFPGPMDWVPRYTGVAPQDQCK
 LTCQARALGYYYVLEPRVVDGTPCSPDSSSVCVQGRCIHAGCDRIIGSKKKFDKCMVCGGDG
 SGCSKQSGSFRKFRYGYNVVTIPAGATHILVRQQGNPGHRSIYLALKLPDGSYALNGEYTL
 MPSPTDVVLPGAVSLRYSGATAASETLSGHGPLAQPLTLQVLVAGNPQDTRLRYSFFVPRPT
 PSTPRPTPDWLHRRRAQILEILRRRPWAGRK

Important features of the protein:

Signal peptide:

amino acids 1-48

N-glycosylation site.

amino acids 68-71

Glycosaminoglycan attachment site

amino acids 188-191, 772-775

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 182-185

Tyrosine kinase phosphorylation site.

amino acids 730-736

N-myristoylation sites.

amino acids 5-10, 19-24, 121-126, 125-130, 130-135, 147-152, 167-
 172, 168-173, 174-179, 323-328, 352-357, 539-544, 555-560, 577-
 582, 679-684, 682-687, 763-768

Amidation sites.

amino acids 560-563, 834-837

Leucine zipper pattern.

amino acids 17-38, 24-45

Neutral zinc metalloproteinases, zinc-binding region signature.

amino acids 358-367

FIGURE 181

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACC**ATGG**
CAAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCC
AAGAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCT
AATTGTCCTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACA
TGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTG
ACCAGAACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTT
TAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGA
TTAAAGTGATTCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACC
ACAACCTTCTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGA
TTTTCTTAAAAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATC
CCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTT
CCTGCCAACGAAAAAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGT
AGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTG
AAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGT
CGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATA
CTGCTACCAAGGAGGACGAGTCATCTGTCTGTGTCATCATGCCTTGTAACCTGGTGGGTGGCCC
GCATGCTGGGGAGGGTCT**TAA**TAGGAGGTTTGAGCTCAAATGCTTAAACTGCTGGCAACATAT
AATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCCTGGTAGCCAGCT
CTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACC
AAAAAAAAAAAAAAAAAAAA

FIGURE 182

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73727

><subunit 1 of 1, 317 aa, 1 stop

><MW: 37130, pI: 5.18, NX(S/T): 3

MAKNPPENCEDCHILNAEAFKSKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAY
DMEHTFYSNGEKKKIYMEIDPVT RTEIFRSGNGTDETLEVHDFKNGYTG IYFVGLQKCFIKT
QIKVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKNSKILEICDNVTMYWI
NPTLISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDY
TENGIEFDPMLDERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWV
ARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 183

GCGGAACTGGCTCCGGCTGGCACCTGAGGAGCGGCGTGACCCCGAGGGCCCAGGGAGCTGCC
 CGGCTGGCCTAGGCAGGCAGCCGCACC**ATG**GCCAGCACGGCCGTGCAGCTTCTGGGCTTCCT
 GCTCAGCTTCCTGGGCATGGTGGGCACGTTGATCACCACCATCCTGCCGCACTGGCGGAGGA
 CAGCGCACGTGGGCACCAACATCCTCACGGCCGTGTCTACCTGAAAGGGCTCTGGATGGAG
 TGTGTGTGGCACAGCACAGGCATCTACCAGTGCCAGATCTACCGATCCCTGCTGGCGCTGCC
 CCAAGACCTCCAGGCTGCCC GCGCCCTCATGGTCATCTCCTGCCTGCTCTCGGGCATAGCCT
 GCGCCTGCGCCGTATCGGGATGAAGTGCACGCGCTGCGCCAAGGGCACACCCGCCAAGACC
 ACCTTTGCCATCCTCGGCGGCACCCTCTTCATCCTGGCCGGCCTCCTGTGCATGGTGGCCGT
 CTCCTGGACCACCAACGACGTGGTGCAGAACTTCTACAACCCGCTGCTGCCCAGCGGCATGA
 AGTTTGAGATTGGCCAGGCCCTGTACCTGGGCTTCATCTCCTCGTCCCTCTCGCTCATTGGT
 GGCACCCTGCTTTGCCTGTCCTGCCAGGACGAGGCACCCTACAGGCCCTACCAGGCCCCGCC
 CAGGGCCACCACGACCACTGCAAACACCGCACCTGCCTACCAGCCACCAGCTGCCTACAAAG
 ACAATCGGGCCCCCTCAGTGACCTCGGCCACGCACAGCGGGTACAGGCTGAACGACTACGTG
TGAGTCCCCACAGCCTGCTTCTCCCCTGGGCTGCTGTGGGCTGGGTCCCCGGCGGGACTGTC
 AATGGAGGCAGGGGTTCAGCACAAAGTTTACTTCTGGGCAATTTTTGTATCCAAGGAAATA
 ATGTGAATGCGAGGAAATGTCTTTAGAGCACAGGGACAGAGGGGGGAAATAAGAGGAGGAGAA
 AGCTCTCTATACCAAAGACTGAAAAAAAAAATCCTGTCTGTTTTTGTATTTATTATATATAT
 TTATGTGGGTGATTTGATAACAAGTTTAATATAAAGTGAAGTTGGGAGTTTGGTCAGTGGGGT
 TGGTTTGTGATCCAGGAATAAACCTTGCGGATGTGGCTGTTTATGAAAAAAAAAAAAA

FIGURE 184

MASTAVQLLGFLLSFLGMVGTLITTLPHWRRTAHVGTNILTAVSYLKGLWMECVWHSTGIY
QCQIYRSLALPQDLQAARALMVISCLLSGIACACAVIGMKCTRCAKGTPAKTTFAILGGTL
FILAGLLCMVAVSWTTNDVVQNFYNPLLPSGMKFEIGQALYLGFISSSLISLIGGTLLCLSCQ
DEAPYRPHYQAPPRATTTTANTAPAYQPPAAYKDNRAPSVTSATHSGYRLNDYV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-103, 115-141, 160-182

FIGURE 185

GAGCTCCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCG
 CAGGCGGCAGGGCGGGCGGCCAGGATC**ATG**TCCACCACCACATGCCAAGTGGTGGCGTTCCT
 CCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCC
 AGGACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGC
 GTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGC
 CATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCC
 TGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCC
 AACATGACACTGACCTCCGGGATCATGTTCAATTGTCTCAGGTCTTTGTGCAATTGCTGGAGT
 GTCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCG
 GCATGGGTGGGATGGTGCAGACTGTTTACAGACCAGGTACACATTTGGTGCGGCTCTGTTCTGTG
 GGCTGGGTGCGCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGG
 CCTGGCACCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTG
 CCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAG
 ATATACGATGGAGGTGCCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTA
 TGTG**TAA**TGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAA
 AACAAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGC
 CTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCC
 ACCATAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTAT
 TTCTTTTTTTTAAATATAACTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCAC
 ATTTTGATGATTTAGACAGACTCCCCCTCTTCCTCCTAGTCAATAAACCCATTGATGATCTA
 TTTCCCAGCTTATCCCCAAGAAAACTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTTT
 CTGCTGTTTGAATTTTGTCTCCCCACCCCCAACTTGGCTAGTAATAAACACTTACTGAAGAA
 GAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTG
 TGATCTTAAAAGTTACCAAACCAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTG
 CTGTTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAG
 TCCTCTTTCTGTGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAAT
 TTAAGTCCTAAATATAGTTAAAATAAATAATGTTTTAGTAAAATGATACACTATCTCTGTGA
 AATAGCCTCACCCCTACATGTGGATAGAAGGAAATGAAAAAATAATTGCTTTGACATTGTCT
 ATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCATGAAAAGCTCACACCTGTAATC
 CTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGTTCGAGACTAGCCTG
 GGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGGTG
 GCATACACCTGTAGTCCCAGCATTCCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGG
 AGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGA
 TCCTGTCTAAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAA
 ACTAATTCTTTAA

FIGURE 186

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73734

><subunit 1 of 1, 261 aa, 1 stop

><MW: 27856, pI: 8.50, NX(S/T): 1

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTE
CRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIM
FIVSGLCAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTL
IGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTE
DEVQSYPSKHDYV

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 187

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCG
 GAGTCCAGCTGGCTAAACTCATCCCAGAGGATA**ATG**GCAACCCATGCCTTAGAAATCGCTG
 GGCTGTTTCTTGGTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGG
 AGAGTGTCGGCCTTCATTGAAAACAACATCGTGGTTTTTGAAGAACTTCTGGGAAGGACTGTG
 GATGAATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGG
 CTCTTTCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTC
 TTGGCTTTCATGATGGCCATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAA
 GGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGC
 TCATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAAT
 GTTGCCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGCT
 GATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTACA
 GATACTCGATACCTTCCCATCGCACAAACCCAAAAAAGTTATCACACCGGAAAGAAGTCACCG
 AGCGTCTACTCCAGAAGTCAGTATGT**G****TAG**TTGTGTATGTTTTTTAACTTTACTATAAAGC
 CATGCAAAATGACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTA
 CTGTTCTTAACTGCCTAATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAA
 GCTATTTTCAGCAGAATGAGATATTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAAT
 TTGTTTTCTAAGGTGGTTCAAGCATCTACTCTTTTTATCATTTACTTCAAAATGACATTGCT
 AAAGACTGCATTATTTTACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAG
 TGTAAACATTTATATCTCACATAGAGACATGCTTATATGGTTTTATTTAAAATGAAATGCCAG
 TCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAGGGTTG
 AAGAAGGTTACTATTAATTGTTTTAAAAACAGCTTAGGGATTAATGTCCCTCCATTTATAATGA
 AGATTAAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTG
 TTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTT
 TTCTTGTTGATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCA
 AACTGCTTTTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATG
 GTTTTAGGAAAGTGAAAATATTTTTGTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAA
 GAAATCATATATGTATGGATATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATC
 AATATAAATAAAAGAGCAGAAAAATATGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACA
 ACAAAAAAAGTTGTCCTTTGAGAACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAAATTG
 TCATTTTTTGTCTGTGAAAAATAAATTTCTTCTGTACCATTTCTGTTTAGTTTTACTAAA
 ATCTGTAAATACTGTATTTTTCTGTTTATTCCAAATTTGATGAACTGACAATCCAATTTGA
 AAGTTTGTGTCGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTTTATACATTTATA
 TTAATAAATTGTACATTTTTTCTAATT

FIGURE 188

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73735

><subunit 1 of 1, 225 aa, 1 stop

><MW: 24845, pI: 9.07, NX(S/T): 0

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRM
QCKIYDSLALLSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGI
IFIITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVF
CCNEKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 189

TCGCC**ATG**GCCTCTGCCGGAATGCAGATCCTGGGAGTCGTCCTGACACTGCTGGGCTGGGTG
AATGGCCTGGTCTCCTGTGCCCTGCCCATGTGGAAGGTGACCGCTTTCATCGGCAACAGCAT
CGTGGTGGCCCAGGTGGTGTGGGAGGGCCTGTGGATGTCCTGCGTGGTGCAGAGCACCGGCC
AGATGCAGTGCAAGGTGTACGACTCACTGCTGGCGCTGCCACAGGACCTGCAGGCTGCACGT
GCCCTCTGTGTCATCGCCCTCCTTGTGGCCCTGTTTCGGCTTGCTGGTCTACCTTGCTGGGGC
CAAGTGTACCACCTGTGTGGAGGAGAAGGATTCCAAGGCCCGCCTGGTGCTCACCTCTGGGA
TTGTCTTTGTCATCTCAGGGGTCCTGACGCTAATCCCCGTGTGCTGGACGGCGCATGCCATC
ATCCGGGACTTCTATAACCCCCCTGGTGGCTGAGGCCCAAAGCGGGAGCTGGGGGCCTCCCT
CTACTTGGGCTGGGCGGCCTCAGGCCTTTTGTGCTGGGTGGGGGGTTGCTGTGCTGCACTT
GCCCCTCGGGGGGGTCCCAGGGCCCCAGCCATTACATGGCCCGCTACTCAACATCTGCCCCT
GCCATCTCTCGGGGGCCCTCTGAGTACCCTACCAAGAATTACGTC**TGA**CGTGGAGGGGAATG
GGGGCTCCGCTGGCGCTAGAGCCATCCAGAAGTGGCAGTGCCCAACAGCTTTGGGATGGGTT
CGTACCTTTTGTCTCTGCCTCCTGCTATTTTTCTTTTGAAGGATATTTAAATTCATTT
GAAACTGAGCCAAGGTGTTGACTCAGACTCTCACTTAGGCTCTGCTGTTTCTCACCCCTTGG
ATGATGGAGCCAAAGAGGGGATGCTTTGAGATTCTGGATCTTGACATGCCCATCTTAGAAGC
CAGTCAAGCTATGGAATAATGCGGAGGCTGCTTGCTGTGCTGGCTTTGCAACAAGACAGAC
TGTCCCCAAGAGTTCCTGCTGCTGCTGGGGGCTGGGCTTCCTTAGATGTCACTGGACAGCTG
CCCCCATCCTACTCAGGTCTCTGGAGCTCCTCTCTTCACCCCTGGAAAAACAAATCATCTG
TTAACAAGGACTGCCCCACCTCCGGAATTCTGACCTCTGTTTCCTCCGTCTGATAAGACG
TCCACCCCCCAGGGCCAGGTCCCAGCTATGTAGACCCCCGCCCCACCTCCAACACTGCACC
CTTCTGCCCTGCCCCCTCGTCTCACCCCTTTACTCACATTTTTATCAAATAAAGCATG
TTTGTAGTGCA

FIGURE 190

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73736
><subunit 1 of 1, 220 aa, 1 stop
><MW: 23292, pI: 8.43, NX(S/T): 0
MASAGMQILGVVLTLLGWVNGLVSCALPMWKVTAFIGNSIVVAQVVWEGWLWMSCVVQSTGQM
QCKVYDSSLALPQDLQAARALCVIALLVALLFGLLVYLAGAKCTTCVEEKDSKARLVLTSGIV
FVISGVLTLIPVCWTAHAIIRDFYNPLVAEAQKRELGASLYLGWAASGLLLLGGGLLCCTCP
SGGSQGPPSHYMARYSTSAPAIRGPPSEYPTKNYV
```

Transmembrane domains:

```
amino acids 8-30 (type II), 82-102, 121-140, 166-186
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220

FIGURE 191

GCCAAGGAGAACATCATCAAAGACTTCTCTAGACTCAAAGGCTTCCACGTTCTACATCTTG
AGCATCTTCTACCACTCCGAATTGAACCAGTCTTCAAAGTAAAGGCAATGGCATTTTATCCC
TTGCAAATTGCTGGGCTGGTTCTTGGGTTCCCTGGCATGGTGGGGACTCTTGCCACAACCCT
TCTGCCTCAGTGGTGGAGTATCAGCTTTTGTGGCAGCAACATTATTGTCTTTGAGAGGCTC
TGGGAAGGGCTCTGG**ATG**AATTGCATCCGACAAGCCAGGGTCCGGTTGCAATGCAAGTTCTA
TAGCTCCTTGTTGGCTCTCCCGCCTGCCCTGGAAACAGCCCCGGGCCCTCATGTGTGTGGCTG
TTGCTCTCTCCTTGATCGCCCTGCTTATTGGCATCTGTGGCATGAAGCAGGTCCAGTGCACA
GGCTCTAACGAGAGGGCCAAAGCATACCTTCTGGGAACCTCAGGAGTCCTCTTCATCCTGAC
GGGTATCTTCGTTCTGATTCCGGTGAGCTGGACAGCCAATATAATCATCAGAGATTTCTACA
ACCCAGCCATCCACATAGGTCAGAAACGAGAGCTGGGAGCAGCACTTTTCCTTGGCTGGGCA
AGCGCTGCTGTCCTCTTCATTGGAGGGGGTCTGCTTTGTGGATTTTGCTGCTGCAACAGAAA
GAAGCAAGGGTACAGATATCCAGTGCCTGGCTACCGTGTGCCACACACAGATAAGCGAAGAA
ATACGACAATGCTTAGTAAGACCTCCACCAGTTATGTC**TAA**TGCCTCCTTTTGGCTCCAAGT
ATGGACTATGGTCAATGTTTTTTATAAAGTCCTGCTAGAAACTGTAAGTATGTGAGGCAGGA
GAACTTGCTTTATGTCTAGATTTACATTGATACGAAAGTTTCAATTTGTTACTGGTGGTAGG
AATGAAAATGACTTACTTGGACATTCTGACTTCAGGTGTATTAAATGCATTGACTATTGTTG
GACCCAATCGCTGCTCCAATTTTCATATTCTAAATTCAAGTATACCCATAATCATTAGCAAG
TGTACAATGATGGACTACTTATTACTTTTTGACCATCATGTATTATCTGATAAGAATCTAAA
GTTGAAATTGATATTCTATAACAATAAAACATATACCTATTCTA

FIGURE 192

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73737
><subunit 1 of 1, 173 aa, 1 stop
><MW: 18938, pI: 9.99, NX(S/T): 1
MNCIRQARVRLQCKFYSSLLALPPALETARALMCVAVALSLIALLLIGICGMKQVQCTGSNER
AKAYLLGTSGVLFILTGIFVLIPVSWTANIIIRDFYNPAIHIGQKRELGAALFLGWASAAVL
FIGGGLLCGFCCCNRRKKQGYRYPVPGYRVPHTDKRRNTTMLSKTSTSYV
```

Important features of the protein:**Transmembrane domains:**

amino acids 31-51, 71-90, 112-133

N-glycosylation site.

amino acids 161-164

FIGURE 193

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGCATGAAGATCACTGGGGGT
CTCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGTCTCC
AAAAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCCCCATCA
CATACTACCAGTTTGTGGTTCTGACTACATCACCTATGGGAATGAATGTCACCTTGTGTACC
GAGAGCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGCTTAAATTCTCCA
TGGACATAGAGAGAAAGGAATGATATTCTCATCATCATCTTCATCATCCCAGGCTCTGACTG
AGTTTCTTTCAGTTTTACTGATGTTCTGGGTGGGGGACAGAGCCAGATTCAGAGTAATCTTG
ACTGAATGGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCAT
TTTTTTTTTAAACACGTCAATAAAAAAATAATCTCCCAGA

196/249

FIGURE 194

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73739
```

><subunit 1 of 1, 85 aa, 1 stop

><MW: 9232, pI: 7.94, NX(S/T): 0

MKITGGLLLLCTVVYFCSSSEAASLSPKKVDCSIYKKYPVVAIPCPITYLPVCGSDYITYGN
ECHLCTESLKSNGRVQFLHDGSC

Signal peptide:

amino acids 1-19

FIGURE 195

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCCCTCGCCCTCCTGCTCGCGCCC
CGCCGCC**ATG**GCTGCCTCCCCGCGCGGCCTGCTGTCCTGGCCCTGACCGGGCTGGCGCTGC
TCCTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAATAAACTCAAGCTGATGCTTCAA
AAACGAGAAGCACCTGTTCCAATAAGACTAAAGTGGCCGTTGATGAGAATAAAGCCAAAGA
ATTCCTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGGACCGGACTCGGCCCCGAGGTGC
AGCAGTGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTTGAAGATGACATC
ACCTATTGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGATTACTACCAACGTCA
CTATGATGAAGACTCTGCAATTGGTCCCCGGAGCCCCCTACGGCTTTAGGCATGGAGCCAGCG
TCAACTACGATGACTAC**TAA**CCATGACTTGCCACACGCTGTACAAGAAGCAAATAGCGATTTC
TCTTCATGTATCTCCTAATGCCTTACACTACTTGGTTTCTGATTTGCTCTATTTTCAGCAGAT
CTTTTCTACCTACTTTGTGTGATCAAAAAAGAAGAGTTAAAACAACACATGTAAATGCCTTT
TGATATTTTCATGGGAATGCCTCTCATTTAAAAATAGAAATAAAGCATTTTGTAAAAAGA

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCCCTCGCCCTCCTGCTCGCGCCC
CGCCGCC**ATG**GCTGCCTCCCCGCGCGGCCTGCTGTCCTGGCCCTGACCGGGCTGGCGCTGC
TCCTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAATAAACTCAAGCTGATGCTTCAA
AAACGAGAAGCACCTGTTCCAATAAGACTAAAGTGGCCGTTGATGAGAATAAAGCCAAAGA
ATTCCTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGGACCGGACTCGGCCCCGAGGTGC
AGCAGTGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTTGAAGATGACATC
ACCTATTGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGATTACTACCAACGTCA
CTATGATGAAGACTCTGCAATTGGTCCCCGGAGCCCCCTACGGCTTTAGGCATGGAGCCAGCG
TCAACTACGATGACTAC**TAA**CCATGACTTGCCACACGCTGTACAAGAAGCAAATAGCGATTTC
TCTTCATGTATCTCCTAATGCCTTACACTACTTGGTTTCTGATTTGCTCTATTTTCAGCAGAT
CTTTTCTACCTACTTTGTGTGATCAAAAAAGAAGAGTTAAAACAACACATGTAAATGCCTTT
TGATATTTTCATGGGAATGCCTCTCATTTAAAAATAGAAATAAAGCATTTTGTAAAAAGA

FIGURE 196

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73742
```

><subunit 1 of 1, 148 aa, 1 stop

><MW: 17183, pI: 8.77, NX(S/T): 0

MAASPARPAVLALTGLALLLLCWGPGGISGNKLKMLQKREAPVPTKTKVAVDENKAKEFL

GSLKRQKRQLWDRTRPEVQQWYQQFLYMGFDEAKFEDDITYWLNDRNGHEY YGDYYQRHYD

EDSAIGPRSPYGFRHGASVNYDDY

Signal peptide:

amino acids 1-30

FIGURE 197

CGGCTCGAGCCCGCCCGGAAGTGCCCGAGGGGCCGCGATGGAGCTGGGGGAGCCGGGCGCTC
GGTAGCGCGGCGGGCAAGGCAGGCGCC**ATG**ACCCTGATTGAAGGGGTGGGTGATGAGGTGAC
CGTCCTTTTCTCGGTGCTTGCCCTGCCTTCTGGTGCTGGCCCTTGCCCTGGGTCTCAACGCACA
CCGCTGAGGGCGGGGACCCACTGCCCCAGCCGTCAGGGACCCCAACGCCATCCCAGCCCAGC
GCAGCCATGGCAGCTACCGACAGCATGAGAGGGGAGGCCCCAGGGGCAGAGACCCCCAGCCT
GAGACACAGAGGTCAAGCTGCACAGCCAGAGCCCAGCACGGGGTTCACAGCAACACCGCCAG
CCCCGGACTCCCCGCAGGAGCCCCTCGTGCTACGGCTGAAATTCCTCAATGATTTCAGAGCAG
GTGGCCAGGGCCTGGCCCCACGACACCATTGGCTCCTTGAAAAGGACCCAGTTTCCCGGCCG
GGAACAGCAGGTGCGACTCATCTACCAAGGGCAGCTGCTAGGCGACGACACCCAGACCCTGG
GCAGCCTTCACCTCCCTCCCAACTGCGTTCTCCACTGCCACGTGTCCACGAGAGTCCGGTCCC
CCAAATCCCCCTGCCCGCCGGGGTCCGAGCCCGGCCCTCCGGGCTGGAAATCGGCAGCCT
GCTGCTGCCCTGCTGCTCCTGCTGTTGCTGCTGCTCTGGTACTGCCAGATCCAGTACCGGC
CCTTCTTTCCCCTGACCGCCACTCTGGGCCTGGCCGGCTTCACCCTGCTCCTCAGTCTCCTG
GCCTTTGCCATGTACCGCCCG**TAG**TGCCCTCCGCGGGCGCTTGGCAGCGTCGCCGGCCCCCTCC
GGACCTTGCTCCCCGCGCCGCGGGGAGCTGCTGCCTGCCCAGGCCCGCCTCTCCGGCCTG
CCTCTTCCCGCTGCCCTGGAGCCCAGCCCTGCGCCGCAGAGGACTCCCGGGACTGGCGGAGG
CCCCGCCCTGCGACCGCCGGGGCTCGGGGCCACCTCCCGGGGCTGCTGAACCTCAGCCCGCA
CTGGGAGTGGGCTCCTCGGGGTGCGGCATCTGCTGTCGCTGCCTCGGCCCCGGGCAGAGCCG
GGCCGCCCCGGGGGCCCGTCTTAGTGTTCTGCCGAGGACCCAGCCGCCTCCAATCCCTGAC
AGCTCCTTGGGCTGAGTTGGGGACGCCAGGTCGGTGGGAGGCTGGTGAAGGGGAGCGGGGAG
GGGCAGAGGAGTTCCCCGGAACCCGTGCAGATTAAAGTAACTGTGAAGTTTTAAAAAAAAA
AAAAAAA

FIGURE 198

MTLIEGVGDEVTVLFSVLACLLVLALAWVSTHTAEGGDPLPQPSGTPTPSQPSAAMAATDSM
RGEAPGAETPSLRHRGQAAQPEPSTGFTATPPAPDSPQEPLVLRLLKFLNDSEQVARAWPHDT
IGSLKRTQFPGREQQVRLIYQGQLLGDDTQTLGSLHLPPNCVLHCHVSTRVGPPNPPCPPGS
EPGPSGLEIGSLLLPLLLLLLLLLLWYCQIQYRPFPLTATLGLAGFTLLLSLLAFAMYRP

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 195-217

FIGURE 199

GAGATTGGAAACAGCCAGGTTGGAGCAGTGAGTGAGTAAGGAAACCTGGCTGCCCTCTCCAG
ATTCCCCAGGCTCTCAGAGAAGATCAGCAGAAAAGTCTGCAAGACCCTAAGAACCATCAGCCC
TCAGCTGCACCTCCTCCCCCTCCAAGGATGACAAAGGCGCTACTCATCTATTTGGTCAGCAGC
TTTCTTGCCCTAAATCAGGCCAGCCTCATCAGTCGCTGTGACTTGGCCCAGGTGCTGCAGCT
GGAGGACTTGGATGGGTTTGAGGGTTACTCCCTGAGTGACTGGCTGTGCCTGGCTTTTGTGG
AAAGCAAGTTCAACATATCAAAGATAAATGAAAATGCGGATGGAAGCTTTGACTATGGCCTC
TTCCAGATCAACAGCCACTACTGGTGCAACGATTATAAGAGTTACTCGGAAAACCTTTGCCA
CGTAGACTGTCAAGATCTGCTGAATCCCAACCTTCTTGCAGGCATCCACTGCGCAAAAAGGA
TTGTGTCCGGAGCACGGGGGATGAACAACCTGGGTAGAATGGAGGTTGCACTGTTTCAGGCCGG
CCACTCTCCTACTGGCTGACAGGATGCCGCCTGAGATTGAAACAGGGTGCGGGTGACCCGTGG
AGTCATTCCAAGACTCCTGTCCTCACTCAGGGATTCTTCATTTCTTCTTCCCTACTGCCTCCA
CTTCATGTTATTTTCTTCCCTTCCCATTACAACTAAACTGACCAGAGCCCCAGGAATAAA
TGGTTTTCTTGGCTTCCTCCTTACTCCCATCTGGACCCAGTCCCCTGGTTCCTGTCTGTTAT
TTGTAAACTGAGGACCACAATAAAGAAATCTTTATATTTATCG

FIGURE 200

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73746

><subunit 1 of 1, 148 aa, 1 stop

><MW: 16896, pI: 6.05, NX(S/T): 1

MTKALLIYLVSSFLALNQASLISRCDLAQVLQLEDLDGFEGYSLSDWLCLAFVESKFNISKI

NENADGSFDYGLFQINSHYWCNDYKSYSENLCNVDCQDLLNPNNLAGIHCAKRIVSGARGMN

NWVEWRLHCSGRPLSYWLTGCRLR

Signal peptide:

amino acids 1-18

203/249

FIGURE 201

TCTGACCTGACTGGAAGCGTCCAAAGAGGGACGGCTGTCAGCCCTGCTTGAAGTGAAGAACCA
 CCAGCTCATCCCAGACACCTCATAGCAACCTATTTATACAAAGGGGAAAGAAACACCTGAG
 CAGAATGGAATCATTATTTTTTTTCCCAAGGAGAAAACCGGGGTAAAGGGAGGGAAGCAATTC
 AATTTGAAGTCCCTGTGAATGGGCTTTTCTGATTTGTAAGTGAAGCAGGTCTTGCACACGC
 TTGGGGTGAACTTGGGTCTGTGGTTTTCTGATTGTAAGTGAAGCAGGTCTTGCACACGC
 TGTGGCAAATGTCAGGACCAGGTTAAGTGAAGTGGCAGAAAACTTCCAGGTGGAACAAGCA
 ACCCATGTTCTGCTGCAAGCTTGAAGGAGCCTGGAGCGGGAGAAAGCTAACTTGAACATGAC
 CTGTTGCATTTGGCAAGTTCTAGCAAC**ATG**CTCCTAAGGAAGCGATACAGGCACAGACCATG
 CAGACTCCAGTTTCTCTGCTGCTCCTGATGCTGGGATGCGTCTGATGATGGTGGCGATGT
 TGCACCCCTCCCCACCACACCTGCAACAGACTGTACAGCCCAAGCCAGCAAGCACAGCCCT
 GAAGCCAGGTACCGCTGGACTTTGGGGAAATCCCAGGATTGGGTACTGGAAGCTGAGGATGA
 GGGTGAAGAGTACAGCCCTCTGGAGGGCCTGCCACCCCTTTATCTCACTGCGGGAGGATCAGC
 TGCTGGTGGCCGTGGCCTTACCCAGGCCAGAAAGGAACCAGAGCCAGGGCAGGAGAGGTGGG
 AGCTACCGCCTCATCAAGCAGCCAAGGAGGCAGGATAAGGAAGCCCCAAAGAGGGACTGGGG
 GGCTGATGAGGACGGGGAGGTGTCTGAAGAAGAGGAGTTGACCCCGTTTACGCTGGACCCAC
 GTGGCCTCCAGGAGGCACTCAGTGCCCGCATCCCCCTCCAGAGGGCTCTGCCCCAGGTGCGG
 CACCCACTGTGTCTGCAGCAGCACCCCTCAGGACAGCCTGCCACAGCCAGCGTCATCCTCTG
 TTTCCATGATGAGGCCTGGTCCACTCTCCTGCGGACTGTACACAGCATCCTCGACACAGTGC
 CCAGGGCCTTCTTGAAGGAGATCATCCTCGTGGACGACCTCAGCCAGCAAGGACAACCTCAAG
 TCTGCTCTCAGCGAATATGTGGCCAGGCTGGAGGGGGTGAAGTTACTCAGGAGCAACAAGAG
 GCTGGGTGCCATCAGGGCCCCGGATGCTGGGGGCCACCAGAGCCACCGGGGATGTGCTCGTCT
 TCATGGATGCCCACTGCGAGTGCCACCCAGGCTGGCTGGAGCCCCCTCCTCAGCAGAATAGCT
 GGTGACAGGAGCCGAGTGGTATCTCCGGTGATAGATGTGATTGACTGGAAGACTTTCCAGTA
 TTACCCCTCAAAGGACCTGCAGCGTGGGGTGTGGACTGGAAGCTGGATTTCCACTGGGAAC
 CTTTGCCAGAGCATGTGAGGAAGGCCCTCCAGTCCCCCATAAGCCCCATCAGGAGCCCTGTG
 GTGCCCGGAGAGGTGGTGGCCATGGACAGACATTACTTCCAAAACACTGGAGCGTATGACTC
 TCTTATGTGCTGCGAGGTGGTGAAAACCTCGAACTGTCTTTCAAGGCCTGGCTCTGTGGTG
 GCTCTGTTGAAATCCTTCCCTGCTCTCGGGTAGGACACATCTACCAAATCAGGATTCCCAT
 TCCCCCTCGACCAGGAGGCCACCCTGAGGAACAGGGTTTCGCATTGCTGAGACCTGGCTGGG
 GTCATTCAAAGAAACCTTCTACAAGCATAGCCCAGAGGCCTTCTCCTTGAGCAAGGCTGAGA
 AGCCAGACTGCATGGAACGCTTGCAGCTGCAAAGGAGACTGGGTGTGCGGACATTCCACTGG
 TTTCTGGCTAATGTCTACCCTGAGCTGTACCCATCTGAACCCAGGCCAGTTTCTCTGGAAA
 GCTCCACAACACTGGAAGTGGGCTCTGTGCAGACTGCCAGGCAGAAGGGGACATCCTGGGCT
 GTCCCATGGTGTGGCTCCTTGCAGTGACAGCCGGCAGCAACAGTACCTGCAGCACACCAGC
 AGGAAGGAGATTCACTTTGGCAGCCCACAGCACCTGTGCTTTGCTGTCAGGCAGGAGCAGGT
 GATTCTTCAGAACTGCACGGAGGAAGGCCTGGCCATCCACCAGCAGCACTGGGACTTCCAGG
 AGAATGGGATGATTGTCCACATTCTTTCTGGGAAATGCATGGAAGCTGTGGTGCAAGAAAAC
 AATAAAGATTTGTACCTGCGTCCGTGTGATGGAAAAGCCCCGCCAGCAGTGGCGATTTGACCA
 GATAAATGCTGTGGATGAACGAT**TGA**ATGTCAATGTGAGAAGGAAAAGAGAATTTTGGCCATC
 AAAATCCAGCTCCAAGTGAACGTAAAGAGCTTATATATTTTCATGAAGCTGATCCTTTTGTGT
 GTGTGCTCCTTGTGTTAGGAGAGAAAAAGCTCTATGAAAGAATATAGGAAGTTTCTCCTTT
 TCACACCTTATTTTATTGACTGCTGGCTGCTTA

FIGURE 202

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73760

><subunit 1 of 1, 639 aa, 1 stop

><MW: 73063, pI: 6.84, NX(S/T): 2

MLLRKRYRHRPCRLQFLLLLLMLGCVLMMVAMLHPPHHTLHQTVTAQASKHSPEARYRLDFG
ESQDWVLEAEDEGEEYSPLEGLPPFISLREDQLLVAVALPQARRNQSQGRRGGSYRLIKQPR
RQDKEAPKRWDGADEDEGEVSEEEELTPFSLDPRGLQEALSARIPLQRALPEVRHPLCLQQHP
QDSLPTASVILCFHDEAWSTLLRTVHSILDTPRAFLKEIILVDDLSQQGQLKSALSEYVAR
LEGVKLLRSNKRLGAIRARMLGATRATGDVLVFMDAHCECHPGWLEPLLSRIAGDRSRVSP
VIDVIDWKTFQYYPSKDLQRGVLDWKLDHFHWEPLPEHVRKALQSPISPIRSPVVPGEVVAMD
RHYFQNTGAYDSLMSLRGGENLELSFKAWLCGGSVEILPCSRVGHYQNDSSHSPLDQEATL
RNRVRIAETWLGSEFKETFYKHSPEAFSLSKAEKPDCEMERLQLQRRRLGCRTFWFLANVYPEL
YPSEPRPSFSGKLHNTGLGLCADCAEGDILGCPMVLAPCSDSRQQQYLQHTSRKEIHFGSP
QHLCFAVRQEQVILQNCTEEGLAIHQQHWDFQENGMIVHILSGKCMEAVVQENNKDLYLRPC
DGKARQQWRFDQINAVDER

Signal peptide:

amino acids 1-28

FIGURE 203

CGCCAAGCATGCGAGTAAAGGCTGAAAATCTGGGTACAGCTGAGGAAGACCTCAGACATGCA
GTCCAGGATGTGGCCTGCGCTGCTGCTGTCCCACCTCCTCCCTCTCTGGCCACTGCTGTTGC
TGCCCCCTCCCACCGCCTGCTCAGGGCTCTTCATCCTCCCCTCGAACCCACCAGCCCCAGCC
CGCCCCCGTGTGCCAGGGGAGGCCCTCGGCCACGTATGTGTGCGTGTGGGAGCGAGC
ACCTCCACCAAGCCGATCTCCTCGGGTCCCAAGATCACGTGCGCAAGTCTGCCTGGCACTG
CACCCCCAGCCACCCCATCAGGCTTTGAGGAGGGGGCCGCCCTCATCCCAATACCCCTGGGCT
ATCGTGTGGGGTCCCACCGTGTCTCGAGAGGATGGAGGGGACCCCAACTCTGCCAATCCCGG
ATTTCTGGACTATGGTTTTTGCAGCCCCCTCATGGGCTCGCAACCCACACCCCAACTCAGACT
CCATGCGAGGTGATGGAGATGGGCTTATCCTTGAGAGGGCACCTGCCACCCTGCGGCCATTCT
CTGTTTGGGGGCGCGTGGGGAAGGTGTGGACCCCCAGCTCTATGTACAAATTACCATCTCCAT
CATCATTTGTTCTCGTGGCCACTGGCATCATCTTCAAGTTCTGCTGGGACCGCAGCCAGAAGC
GACGCAGACCCTCAGGGCAGCAAGGTGCCCTGAGGCAGGAGGAGAGCCAGCAGCCACTGACA
GACCTGTCCCCGGCTGGAGTCACTGTGCTGGGGGCCTTCGGGGACTCACCTACCCCCACCC
TGACCATGAGGAGCCCCGAGGGGGACCCCGGCCTGGGATGCCCCACCCCAAGGGGGCTCCAG
CCTTCCAGTTGAACCGTGAAGGGCAGGGGCAATGGGATGGGAGGGCAAAGAGGGAAGGCAAC
TTAGGTCTTCAGAGCTGGGGTGGGGTGCCCTCTGGATGGGTAGTGAGGAGGCAGGCGTGGC
CTCCACAGCCCCCTGGCCCTCCAAGGGGGCTTGACCAGCTCCTCTCTGGGAGGCACCCTTC
CTTCTCCAGTCTCTCAGGATCTGTGTCTTATTCTCTGCTGCCATAACTCCAACCTCTGCC
TCTTTGGTTTTTTTTCTCATGCCACCTTGTCTAAGACAACCTCTGCCCTCTTAACCTTGATTCC
CCTCTTTGTCTTGAACCTTCCCTTCTATTCTGGCCTACCCCTTGTTCTGACTGTGCCCTT
TCCCTCTTCTCTCAGGATTCCCTTGGTGAATCTGTGATGCCCCAATGTTGGGGTGCAGCC
AAGCAGGAGGCCAAGGGGGCCGGCACAGCCCCCATCCCACCTGAGGGTGGGGCAGCTGTGGGGA
GCTGGGGCCACAGGGGCTCCTGGCTCCTGCCCTTGACACACCACCCGGAACACTCCCCAGCC
CCACGGGCAATCCTATCTGCTCGCCCTCCTGCAGGTGGGGGCCTCACATATCTGTGACTTCG
GGTCCCTGTCCCCACCCTTGTGCACTCACATGAAAGCCTTGACACACTCACCTCCACCTTCAC
AGGCCATTTGCACACGCTCCTGCACCTCTCCCCGTCCATACCGCTCCGCTCAGCTGACTCT
CATGTTCTCTCGTCTCACATTTGCACTCTCTCCTTCCCACATTCTGTGCTCAGCTCACTCAG
TGGTCAGCGTTTTCTGCACACTTTACCTCTCATGTGCGTTTTCCCGGCCTGATGTTGTGGTGG
TGTGCGGCGTGTCACTCTCTCCCTCATGAACACCCACCCACCTCGTTTTCCGCAGCCCCCTGC
GTGCTGCTCCAGAGGTGGGTGGGAGGTGAGCTGGGGGCTCCTTGGGCCCTCATCGGTCATGG
TCTCGTCCCATTCACACCATTTGTTTTCTGTCTCCCCATCCTACTCCAAGGATGCCGGCA
TCACCCTGAGGGCTCCCCCTTGGGAATGGGGTAGTGAGGGCCCCAGACTTCACCCCCAGCCCA
CTGCTAAAATCTGTTTTCTGACAGATGGGTTTTTGGGGAGTCGCCCTGCTGCACTACATGAGAA
AGGGACTCCCATTTGCCCTTCCCTTTCTCCTACAGTCCCTTTTTGTCTTGTCTGTCTGGCTG
TCTGTGTGTGTGCCATTCTCTGGACTTCAGAGCCCCCTGAGCCAGTCTCCTCCTTCCCAGCCT
CCCTTTGGGCCTCCCTAACTCCACCTAGGCTGCCAGGGACCGGAGTCAGCTGGTTCAAGGCC
ATCGGGAGCTCTGCCTCCAAGTCTACCCTTCCCTTCCCGGACTCCCTCCTGTCCCCCTCCTTT
CCTCCCTCCTTCCCTTCCACTCTCCTTCCCTTTTGTCTCCCTGCCCTTCCCCCTCCTCAGGT
CTTCCCTCCTTCTCACTGGTTTTTCCACCTTCCCTCCTTCCCTTCTTCCCTGGCTCCTAGGCT
GTGATATATATTTTTGTATTATCTCTTCTTCTTCTTGTGGTGATCATCTTGAATTACTGTG
GGATGTAAGTTTTCAAATTTTTCAAATAAAGCCTTTGCAAGATAA

FIGURE 204

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393

><subunit 1 of 1, 243 aa, 1 stop

><MW: 26266, pI: 8.43, NX(S/T): 1

MRPQGPAASPQRLRGLLLLLLLLQLPAPSSASEIPKGGKQKAQLRQREVVDLYNGMCLQGPAGV
PGRDGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMN
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 205

GTTAACCAGCGCAGTCCTCCGTGCGTCCCGCCCGCCGCTGCCCTCACTCCCGGCCAGG**ATGG**
CATCCTGTCTGGCCCTGCGCATGGCGCTGCTGCTGGTCTCCGGGGTTCTGGCCCCCTGCGGTG
CTCACAGACGATGTTCCACAGGAGCCCGTGCCACGCTGTGGAACGAGCCGGCCGAGCTGCC
GTCGGGAGAAGGCCCCGTGGAGAGCACCCAGCCCCGGCCGGGAGCCCGTGGACACCGGTCCCC
CAGCCCCCACCCTCGCGCCAGGACCCGAGGACAGCACCGCGCAGGAGCGGCTGGACCAGGGC
GGCGGGTTCGCTGGGGCCCCGGCGCTATCGCGGCCATCGTGATCGCCGCCCTGCTGGCCACCTG
CGTGGTGCTGGCGCTCGTGGTCGTCGCGCTGAGAAAGTTTTCTGCCTCC**TGA**AGCGAATAAA
GGGGCCGCGCCCGGCCGCGGCGCGACTCGGCAAAAAAAAAAAAAA

FIGURE 206

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76398
><subunit 1 of 1, 121 aa, 1 stop
><MW: 12073, pI: 4.11, NX(S/T): 0
MASCLALRMALLLVSGVLAPAVLTDDVPQEPVPTLWNEPAELPSGEGPVVESTSPGREPVDTG
PPAPTVPAGPEDSTAQERLDQGGGSLGPGAIAAIVIAALLATCVVLALVVVALRKFSAS
```

Important features of the protein:**Signal peptide:**

amino acids 1-19

Transmembrane domain:

amino acids 91-110

Glycosaminoglycan attachment site.

amino acids 44-47

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 116-119

N-myristoylation site.

amino acids 91-96

FIGURE 207

GGCCGTTGGTTGGTGCGCGGCTGAAGGGTGTGGCGCGAGCAGCGTCGTTGGTTGGCCGGCGG
CGGGCCGGGACGGGC**ATG**GCCCTGCTGCTGTGCCTGGTGTGCCTGACGGCGGCGCTGGCCCA
CGGCTGTCTGCACTGCCACAGCAACTTCTCCAAGAAGTTCTCCTTCTACCGCCACCATGTGA
ACTTCAAGTCCTGGTGGGTGGGCGACATCCCCGTGTCAGGGGCGCTGCTCACCGACTGGAGC
GACGACACGATGAAGGAGCTGCACCTGGCCATCCCCGCCAAGATCACCCGGGAGAAGCTGGA
CCAAGTGGCGACAGCAGTGTACCAGATGATGGATCAGCTGTACCAGGGGAAGATGTACTTCC
CCGGGTATTTCCCCAACGAGCTGCGAAACATCTTCCGGGAGCAGGTGCACCTCATCCAGAAC
GCCATCATCGAAAGGCACCTGGCACCAGGCAGCTGGGGAGGAGGGCAGCTCTCCAGGGAGGG
ACCCAGCCTAGCACCTGAAGGATCAATGCCATCACCCCGCGGGGACCTCCCC**TAA**GTAGCCC
CCAGAGGCGCTGGGAGTGTTGCCACCGCCCTCCCCTGAAGTTTGCTCCATCTCACGCTGGGG
GTCAACCTGGGGACCCCTTCCCTCCGGGCCATGGACACACATACATGAAAACCAGGCCGCAT
CGACTGTCAGCACCGCTGTGGCATCTTCCAGTACGAGACCATCTCCTGCAACAACCTGCACAG
ACTCGCACGTGCGCTGCTTTGGCTATAACTGCGAGTAGGGCTCAGGCATCACACCCACCCGT
GCCAGGGCCCTACTGTCCCTGGGGTCCCAGGCTCTCCTTGGAGGGGGCTCCCCGCCTTCCAC
CTGGCTGTTCATCGGGTAGGGCGGGGCCGTGGGTTCAGGGGCGCACCACTTCCAAGCCTGTGT
CCCACAGGTCCTCGGCGCAGTGGAAGTCAGCTGTCCAGGGCCTCCTGAACTACATAAATAAC
TGGCACAAGTAAGTCCCCCTCCTCAAACCAACACAGGCAGTGTGTGTATGTGAGCACCTCGTG
GGTGAGTATGTGTGGGGCACAGGCTGGCTCCCTCAGCTCCCACGTCCTAGAGGGGGCTCCCGA
GGAGGTGGAACCTCAACCCAGCTCTGCGCAGGAGGCGGCTGCAGTCCTTTTCTCCCTCAAAG
GTCTCCGACCCTCAGCTGGAGGCGGGCATCTTTCCTAAAGGGTCCCCATAGGGTCTGGTTCC
ACCCCATCCCAGGTCTGTGGTCAGAGCCTGGGAGGGTTCCCTACGATGGTTAGGGGTGCCCC
ATGGAGGGGCTGACTGCCCCACATTGCCTTTTCAGACAGGACACGAGCATGAGGTAAGGCCGC
CCTGACCTGGACTTCAGGGGGAGGGGGTAAAGGGAGAGAGGAGGGGGGCTAGGGGGTCTCT
AGATCAGTGGGGGCACTGCAGGTGGGGCTCTCCCTATACCTGGGACACCTGCTGGATGTCAC
CTCTGCAACCACACCCATGTGGTGGTTTCATGAACAGACCACGCTCCTCTGCCTTCTCCTGG
CCTGGGACACACAGAGCCACCCCGGCCTTGTGAGTGACCCAGAGAAGGGAGGCCTCGGGAGA
AGGGGTGCTCGTAAGCCAACACCAGCGTGCCGCGGCCTGCACACCCTTCGGACATCCCAGGC
ACGAGGGTGTCTGGATGTGGCCACACATAGGACCACACGTCCCAGCTGGGAGGAGAGGCCT
GGGGCCCCCAGGGAGGGAGGCAGGGGGTGGGGGACATGGAGAGCTGAGGCAGCCTCGTCTCC
CCGCAGCCTGGTATCGCCAGCCTTAAGGTGTCTGGAGCCCCCACACTTGGCCAACCTGACCT
TGGAAGATGCTGCTGAGTGTCTCAAGCAGCACTGACAGCAGCTGGGCCTGCCCCAGGGCAAC
GTGGGGGCGGAGACTCAGCTGGACAGCCCCTGCCTGTCACTCTGGAGCTGGGCTGCTGCTGC
CTCAGGACCCCCCTCTCCGACCCCGGACAGAGCTGAGCTGGCCAGGGCCAGGAGGGCGGGAGG
GAGGGAATGGGGGTGGGCTGTGCGCAGCATCAGCGCCTGGGCAGGTCCGCAGAGCTGCGGGA
TGTGATTAAAGTCCCTGATGTTTCTC

FIGURE 208

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76399

><subunit 1 of 1, 157 aa, 1 stop

><MW: 17681, pI: 7.65, NX(S/T): 1

MALLLCLVCLTAALAHGCLHCHSNFSKKFSFYRHHVNFKSWWVGDI PVSGALLTDWSDDTMK
ELHLAIPAKITREKLDQVATAVYQMMDQLYQGKMYFPGYFPNELRNIFREQVHLIQNAIIER
HLAPGSWGGGQLSREGPSLAPEGSMPSPRGDLP

Signal peptide:

amino acids 1-15

20140101 14:00:00

FIGURE 209

AGCAGGAGCAGGAGAGGGACA**ATG**GAAAGCTGCCCCGTCCAGGTTTCATGTTTCCTCTTATTTCT
 CCTCACGTGTGAGCTGGCTGCAGAAAGTTGCTGCAGAAAGTTGAGAAATCCTCAGATGGTCCTG
 GTGCTGCCCAGGAACCCACGTGGCTCACAGATGTCCCAGCTGCCATGGAATTCATTGCTGCC
 ACTGAGGTGGCTGTCATAGGCTTCTTCCAGGATTTAGAAATACCAGCAGTGCCCATACTCCA
 TAGCATGGTGCAAAAATTCCCAGGCGTGTCAATTTGGGATCAGCACTGATTCTGAGGTTCTGA
 CACACTACAACATCACTGGGAACACCATCTGCCTCTTTCGCCTGGTAGACAATGAACAACCTG
 AATTTAGAGGACGAAGACATTGAAAGCATTGATGCCACCAAATTGAGCCGTTTCATTGAGAT
 CAACAGCCTCCACATGGTGACAGAGTACAACCCTGTGACTGTGATTGGGTTATTCAACAGCG
 TAATTCAGATTCATCTCCTCCTGATAATGAACAAGGCCTCCCCAGAGTATGAAGAGAACATG
 CACAGATACCAGAAGGCAGCCAAGCTCTTCCAGGGGAAGATTCTCTTTATTCTGGTGGACAG
 TGGTATGAAAGAAAATGGGAAGGTGATATCATTTTTCAAATAAGGAGTCTCAACTGCCAG
 CTTTGGCAATTTACCAGACTCTAGATGACGAGTGGGATACACTGCCACAGCAGAAGTTTCC
 GTAGAGCATGTGCAAACTTTTGTGATGGATTCTTAAGTGGAAAATTGTTGAAAGAAAATCG
 TGAATCAGAAGGAAAGACTCCAAAGGTGGAACCT**TGA**CTTCTCCTTGGAACCTACATATGGCC
 AAGTATCTACTTTTATGCAAAGTAAAAAGGCACAACCTCAAATCTCAGAGACACTAAACAACAG
 GATCACTAGGCCTGCCAACCACACACACACGCACGTGCACACACGCACGCACGCGTGCACAC
 ACACACGCGCACACACACACACACACAGAGCTTCATTTCCCTGTCTTAAATCTCGTTTTCTC
 TTCTTCCTTCTTTTAAATTTTCATATCCTCACTCCCTATCCAATTTCTTCTTATCGTGCATT
 CATACTCTGTAAGCCCATCTGTAACACACCTAGATCAAGGCTTTAAGAGACTCACTGTGATG
 CCTCTATGAAAGAGAGGCATTCCCTAGAGAAAGATTGTTCCAATTTGTCATTTAATATCAAGT
 TTGTATACTGCACATGACTTACACACAACATAGTTCCCTGCTCTTTTAAGGTTACCTAAGGGT
 TGAAACTCTACCTTCTTTCATAAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCAAAGG
 ATGGTTTTTAAACACCTTTGTGAAATTGTCTTTTGGCAGAAGTTAAAGGCTGTCTCCAAGTC
 CCTGAACTCAGCAGAAATAGACCATGTGAAAACCTCATGCTTGGTTAGCATCTCCAACCTCCC
 TATGTAAATCAACAACCTGCATAATAAATAAAAGGCAATCATGTTATA

FIGURE 210

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76401

><subunit 1 of 1, 273 aa, 1 stop

><MW: 30480, pI: 4.60, NX(S/T): 1

MEAAPSRFMFLFLLTCELAAEVAAEVEKSSDGPAAQEPTWLTDVPAAMEFIAATEVAVIG
FFQDLEIPAVPILHSMVQKFPGVSGISTDSEVLTHYNITGNTICLFRLLVDNEQLNLEDEDI
ESIDATKLSRFIEINSLHMTVEYNPVTVIGLFNSVVIQIHLILLIMNKASPEYEENMHRYQKAA
KLFQGGKILFILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDDEWDTLPTAEVSVEHVQNF
CDGFLSGKLLKENRESEGKTPKVEL

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 143-162

FIGURE 211

GGAGAGCCGCGGCTGGGACCGGAGTGGGGAGCGCGGCGTGGAGGTGCCACCCGGCGCGGGTG
GCGGAGAGATCAGAAGCCTCTTCCCCAAGCCGAGCCAACCTCAGCGGGGACCCGGGCTCAGG
GACGCGGCGGCGGCGGCGGCGGCGACTGCAGTGGCTGGACG**ATG**GCAGCGTCCGCCGGAGCCGGG
GCGGTGATTGCAGCCCCAGACAGCCGGCGCTGGCTGTGGTGGTGGTGGTGGCGGCGGCGCTTGG
GCTCTTGACAGCTGGAGTATCAGCCTTGGAAGTATATACGCCAAAAGAAATCTTCGTGGCAA
ATGGTACACAAGGGAAGCTGACCTGCAAGTTCAAGTCTACTAGTACGACTGGCGGGTTGACC
TCAGTCTCCTGGAGCTTCCAGCCAGAGGGGGCCGACACTACTGTGTCGTTTTTCCACTACTC
CCAAGGGCAAGTGTACCTTGGGAATTATCCACCATTTAAAGACAGAATCAGCTGGGCTGGAG
ACCTTGACAAGAAAGATGCATCAATCAACATAGAAAATATGCAGTTTATACACAATGGCACC
TATATCTGTGATGTCAAAAACCCTCCTGACATCGTTGTCCAGCCTGGACACATTAGGCTCTA
TGTCGTAGAAAAAGAGAATTTGCCTGTGTTTCCAGTTTGGGTAGTGGTGGGCATAGTTACTG
CTGTGGTCCTAGGTCTCACTCTGCTCATCAGCATGATTCTGGCTGTCCTCTATAGAAGGAAA
AACTCTAAACGGGATTACACTGGCTGCAGTACATCAGAGAGTTTGTCCACCAGTTAAGCAGGC
TCCTCGGAAGTCCCCCTCCGACACTGAGGGTCTTGTAAGAGTCTGCCTTCTGGATCTCACC
AGGGCCCAGTCATATATGCACAGTTAGACCACTCCGGCGGACATCACAGTGACAAGATTAAC
AAGTCAGAGTCTGTGGTGTATGCGGATATCCGAAAGAAT**TAA**GAGAATACCTAGAACATATC
CTCAGCAAGAAACAAAACCAAACTGGACTCTCGTGCAGAAAATGTAGCCCATTACCACATGT
AGCCTTGGAGACCCAGGCAAGGACAAGTACACGTGTACTCACAGAGGGAGAGAAAGATGTGT
ACAAAGGATATGTATAAATATTCTATTTAGTTCATCCTGATATGAGGAGCCAGTGTTGCATGA
TGAAAAGATGGTATGATTCTACATATGTACCCATTGTCTTGCTGTTTTTGTACTTTCTTTTC
AGGTCATTTACAATTGGGAGATTTAGAAAACATTCCTTTACCATCATTTAGAAATGGTTTG
CCTTAATGGAGACAATAGCAGATCCTGTAGTATTTCCAGTAGACATGGCCTTTTAATCTAAG
GGCTTAAGACTGATTAGTCTTAGCATTTACTGTAGTTGGAGGATGGAGATGCTATGATGGAA
GCATACCCAGGGTGGCCTTTAGCACAGTATCAGTACCATTTATTTGTCTGCCGCTTTTAAAA
AATACCCATTGGCTATGCCACTTGAAAACAATTTGAGAAGTTTTTTTGAAGTTTTTCTCACT
AAAATATGGGGCAATTGTTAGCCTTACATGTTGTGTAGACTTACTTTAAGTTTGCACCCTTG
AAATGTGTCATATCAATTTCTGGATTACATAATAGCAAGATTAGCAAAGGATAAATGCCGAAG
GTCACTTCATTCTGGACACAGTTGGATCAATACTGATTAAGTAGAAAATCCAAGCTTTGCTT
GAGAACTTTTGTAACGTGGAGAGTAAAAAGTATCGGTTTTA

FIGURE 212

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76510

><subunit 1 of 1, 269 aa, 1 stop

><MW: 29082, pI: 9.02, NX(S/T): 3

MAASAGAGAVIAAPDSRRWLWSVLAAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKFKS
TSTTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDDKSDASINIEN
MQFIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVVGIVTAVVLGLTLLISMI
LAVLYRRKNSKRDTGCTSESLSPVKQAPRKSPSDTEGLVKSLPSGSHQGPVIYAQLDHSG
GHHSDKINKSESVVYADIRKN

Signal peptide:

amino acids 1-37

Transmembrane domain:

amino acids 161-183

FIGURE 213

GCCGGCTGTGCAGAGACGCC**ATG**TACCGGCTCCTGTCAGCAGTGACTGCCCCGGGCTGCCGCC
CCCCGGGGGCTTGGCCTCAAGCTGCGGACGACGCGGGGTCCATCAGCGCGCCGGGCTGCCGCC
TCTCGGCCACGGCTGGGTGCGGGGCCTCGGGCTGGGGCTGGGGCTGGCGCTCGGGGTGAAGC
TGGCAGGTGGGCTGAGGGGCGCGGCCCCGCGCAGTCCCCCGCGGCCCCCGACCCTGAGGCG
TCGCCTCTGGCCGAGCCGCCACAGGAGCAGTCCCTCGCCCCGTGGTCTCCGCAGACCCCGGC
GCCGCCCTGCTCCAGGTGCTTCGCCAGAGCCATCGAGAGCAGCCGCGACCTGCTGCACAGGA
TCAAGGATGAGGTGGGCGCACCGGGCATAGTGGTTGGAGTTTCTGTAGATGGAAAAGAAGTC
TGGTCAGAAGGTTTAGGTTATGCTGATGTTGAGAACCGTGTACCATGTAAACCAGAGACAGT
TATGCGAATTGCTAGCATCAGCAAAAGTCTCACCATGGTTGCTCTTGCCAAATTGTGGGAAG
CAGGGAAACTGGATCTTGATATTCCAGTACAACATTATGTTCCCGAATTCCCAGAAAAAGAA
TATGAAGGTGAAAAGGTTTCTGTACACAAGATTACTGATTTCCCATTTAAGTGAATTTCG
TCATTATGAAAAGGACATAAAAAAGGTGAAAGAAGAGAAAGCTTATAAAGCCTTGAAGATGA
TGAAAGAGAATGTTGCATTTGAGCAAGAAAAAGAAGGCAAAGTAATGAAAAGAATGATTTT
ACTAAATTTAAAACAGAGCAGGAGAATGAAGCCAAATGCCGGAATTCAAACCTGGCAAGAA
AAAGAATGATTTTGAACAAGGCGAATTATATTTGAGAGAAAAGTTTGAAAATTCAATTGAAT
CCCTAAGATTATTTAAAAATGATCCTTTGTTCTTCAAACCTGGTAGTCAGTTTTTGTATTCA
ACTTTTGGCTATACCCTACTGGCAGCCATAGTAGAGAGAGCTTCAGGATGTAAATATTTGGA
CTATATGCAGAAAATATTCCATGACTTGGATATGCTGACGACTGTGCAGGAAGAAAACGAGC
CAGTGATTTACAATAGAGCAAGG**TAA**ATGAATACCTTCTGCTGTGTCTAGCTATATCGCATC
TTAACACTATTTTATTAATTAAGTCAAATTTTCTTTGTTTCCATTCCAAAATCAACCTGC
CACATTTTGGGAGCTTTTCTACATGTCTGTTTTCTCATCTGTAAAGTGAAGGAAGTAAACA
TGTTTATAAAGTAAAAAAA

FIGURE 214

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76522
><subunit 1 of 1, 373 aa, 1 stop
><MW: 41221, pI: 8.54, NX(S/T): 0
MYRLLSAVTARAAAPGGLASSCGRRGVHQRAGLPPLGHGWVGGLGLGLGLALGVKLAGGLRG
AAPAQSPAAPDPEASPLAEPPEQESLAPWSPQTPAPPCSRCFARAIESSRDLLHRIKDEVGA
PGIVVGVSVDGKEVWSEGLGYADVENRVPCKPETVMRIASISKSLTMVALAKLWEAGKLDLD
IPVQHYVPEFPEKEYEGEKVSVTTRLLISHLSGIRHYEKDIKKVKEEKAYKALKMMKENVAF
EQEKEGKSNEKNDFTKFKTEQENEAKCRNSKPGKKKNDFEQGELYLREKFENSIESLRLFKN
DPLFFKPGSQFLYSTFGYTLLAAIVERASGCKYLDYMQKIFHDLDMLTTVQEENEPVIYNRAR
```

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 39-60

FIGURE 215

GTGACACTATAGAAGAGCTATGACGTCGCATGCACGCGTACGTAAGCTCGGAATTCGGCTCG
AGGCTGGTGGGAAGAAGCCGAG**ATG**GCGGCAGCCAGCGCTGGGGCAACCCGGCTGCTCCTGC
TCTTGCTGATGGCGGTAGCAGCGCCCAGTCGAGCCCGGGGCAGCGGCTGCCGGGCCGGGACT
GGTGCGCGAGGGGCTGGGGCGGAAGGTGAGAGGGCGAGGCCTGTGGCACGGTGGGGCTGCT
GCTGGAGCACTCATTTGAGATCGATGACAGTGCCAACTTCCGGAAGCGGGGCTCACTGCTCT
GGAACCAGCAGGATGGTACCTTGTCCCTGTCACAGCGGCAGCTCAGCGAGGAGGAGCGGGGC
CGACTCCGGGATGTGGCAGCCCTGAATGGCCTGTACCGGGTCCGGATCCCAAGGCGACCCGG
GGCCCTGGATGGCCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTCCCTGCGTGCTCCCTGG
TGGAGTCGCACCTGTCGGACCAGCTGACCCTGCACGTGGATGTGGCCGGCAACGTGGTGGGC
GTGTCGGTGGTGACGCACCCCGGGGGCTGCCGGGGGCCATGAGGTGGAGGACGTGGACCTGGA
GCTGTTCAACACCTCGGTGCAGCTGCAGCCGCCCACCACAGCCCCAGGCCCTGAGACGGCGG
CCTTCATTGAGCGCCTGGAGATGGAACAGGCCCAAGAAGGCCAAGAACCCCGAGGAGCAGAAG
TCCTTCTTCGCCAAATACTGGATGTACATCATTCCCGTCGTCCTGTTCCCTCATGATGTCAGG
AGCGCCAGACACCGGGGGCCAGGGTGGGGGTGGGGGTGGGGGTGGTGGTGGGGGTAGTGGCC
TTTGCTGTGTGCCACCCTCCCTG**TAA**GTCTATTTAAAAACATCGACGATACATTGAAATGTG
TGAACGTTTTGAAAAGCTACAGCTTCCAGCAGCCAAAAGCAACTGTTGTTTTGGCAAGACGG
TCCTGATGTACAAGCTTGATTGAAATTCAGTGTCACTTGATACGTTATTCAGAAACCCAAG
GAATGGCTGTCCCCATCCTCATGTGGCTGTGTGGAGCTCAGCTGTGTTGTGTGGCAGTTTAT
TAAACTGTCCCCCAGATCGACACGCAAAAAAAAAA

FIGURE 216

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76529
><subunit 1 of 1, 269 aa, 1 stop
><MW: 28004, pI: 5.80, NX(S/T): 1
MAAASAGATRLLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGREGEACGTVGLLLEHSFEI
DDSANFRKRGSLLWNQQDGTLSLSQRQLSEEERGRLRDVAALNGLYRVRIPRRPGALDGLEA
GGYVSSFVPACSLVESHLSDDLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLELFNTSVQ
LQPPTTAPGPETAAFIERLEMEQAQKAKNPQEQKSFFAKYWMYIIPVVLFLMMSGAPDTGGQ
GGGGGGGGGGGGSGGLCCVPPSL
```

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 226-243

FIGURE 217

GGAGCGCTGCTGGAACCCGAGCCGGAGCCGGAGCCACAGCGGGGAGGGTGGCCTGGCGGCCT
GGAGCCGGACGTGTCCGGGGCGTCCCCGAGACCGGGGAGCAGGTCTCCGGGGGGCCCACC
ATGCTGGTGACTGCCTACCTTGCTTTTGTAGGCCTCCTGGCCTCCTGCCTGGGGCTGGAAC
GTCAAGATGCCGGGCTAAACCCCTGGAAGGGCCTGCAGCAATCCCTCCTTCCTTCGGTTTC
AACTGGACTTCTATCAGGTCTACTTCCTGGCCCTGGCAGCTGATTGGCTTCAGGCCCCCTAC
CTCTATAAACTCTACCAGCATTACTACTTCCTGGAAGGTCAAATTGCCATCCTCTATGTCTG
TGGCCTTGCCTCTACAGTCCTCTTTGGCCTAGTGGCCTCCTCCCTTGTGGATTGGCTGGGTC
GCAAGAATTCTTGTGTCTCTTCTCCCTGACTTACTCACTATGCTGCTTAACCAAACCTCTCT
CAAGACTACTTTGTGCTGCTAGTGGGGCGAGCACTTGGTGGGCTGTCCACAGCCCTGCTCTT
CTCAGCCTTCGAGGCCTGGTATATCCATGAGCACGTGGAACGGCATGACTTCCCTGCTGAGT
GGATCCCAGCTACCTTTGCTCGAGCTGCCTTCTGGAACCATGTGCTGGCTGTAGTGGCAGGT
GTGGCAGCTGAGGCTGTAGCCAGCTGGATAGGGCTGGGGCCTGTAGCGCCCTTTGTGGCTGC
CATCCCTCTCCTGGCTCTGGCAGGGGCCTTGGCCCTTCGAAACTGGGGGGAGAACTATGACC
GGCAGCGTGCCTTCTCAAGGACCTGTGCTGGAGGCCTGCGCTGCCTCCTGTGCGACCGCCGC
GTGCTGCTGCTGGGCACCATAACAAGCTCTATTTGAGAGTGTGTCATCTTCATCTTTGTCTTCCT
CTGGACACCTGTGCTGGACCCACACGGGGGCCCTCTGGGCATTATCTTCTCCAGCTTCATGG
CAGCCAGCCTGCTTGGCTCTTCCCTGTACCGTATCGCCACCTCCAAGAGGTACCACCTTCAG
CCCATGCACCTGCTGTCCCTTGCTGTGCTCATCGTCGTCTTCTCTCTCTTCATGTTGACTTT
CTCTACCAGCCCAGGCCAGGAGAGTCCGGTGGAGTCCTTCATAGCCTTTCTACTTATTGAGT
TGGCTTGTGGATTATACTTTCCAGCATGAGCTTCCTACGGAGAAAGGTGATCCCTGAGACA
GAGCAGGCTGGTGTACTCAACTGGTTCCGGGTACCTCTGCACTCACTGGCTTGCCTAGGGCT
CCTTGTCTCCATGACAGTGATCGAAAAACAGGCACTCGGAATATGTTTCAGCATTTGCTCTG
CTGTCATGGTGATGGCTCTGCTGGCAGTGGTGGGACTCTTCACCGTGGTAAGGCATGATGCT
GAGCTGCGGGTACCTTCACCTACTGAGGAGCCCTATGCCCCCTGAGCTG**TAA**CCCCACTCCAG
GACAAGATAGCTGGGACAGACTCTTGAATTCCAGCTATCCGGGATTGTACAGATCTCTCTGT
GACTGACTTTTGTGACTGTCCTGTGGTTTCTCCTGCCATTGCTTTGTGTTTGGGAGGACATGA
TGGGGGTGATGGACTGGAAAGAAGGTGCCAAAAGTTCCCTCTGTGTTACTCCCATTTAGAAA
ATAAACACTTTTAAATGATCAAAAAAAAAAAAA

FIGURE 218

MLVTAYLAFVGLLASCLGLELSRCRAKPPGRACSNPSFLRFQLDFYQVYFLALAADWLQAPY
LYKLYQHYYFLEGQIAILYVCGLASTVLFGLVASSLVDWLGRKNSCVLFSLTYSLCCLTKLS
QDYFVLLVGRALGGLSTALLFSAFEAWYIHEHVERHDFPAEWIPATFARAAFWNHVLAVVAG
VAAEAVASWIGLGPVAPFVAAIPLLALAGALALRNWGENYDRQRAFSRTCAGGLRCLLSDRR
VLLLGTIQALFESVIFIFVFLWTPVLDPHGAPLGIIFFSSFMAASLLGSSLYRIATSKRYHLQ
PMHLLSLAVLIVVFSLFMLTFSTSPGQESPVESFIAFLLLIELACGLYFPSMSFLRRKVIPET
EQAGVLNWFVRVPLHSLACLGLLVLHDSRKTGTRNMFISCSAVMMALLAVVGLFTVVRHDA
ELRVPSPTTEEPYAPEL

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 41-55, 75-94, 127-143, 191-213, 249-270, 278-299,
314-330, 343-359, 379-394, 410-430

FIGURE 219

GCGACGCGCGGCGGGGCGGCGAGAGGAAACGCGGCGCCGGGCGGGGCCCGGCCCTGGAG**ATG**
GTCCCCGGCGCCGCGGGCTGGTGTGTCTCGTGCTCTGGCTCCCCGCGTGCGTCGCGGCCCA
CGGCTTCCGTATCCATGATTATTTGTACTTTCAAGTGCTGAGTCCTGGGGACATTCGATACA
TCTTCACAGCCACACCTGCCAAGGACTTTGGTGGTATCTTTCACACAAGGTATGAGCAGATT
CACCTTGTCCCCGCTGAACCTCCAGAGGCCTGCGGGGAACTCAGCAACGGTTTCTTCATCCA
GGACCAGATTGCTCTGGTGGAGAGGGGGGGCTGCTCCTTCCTCTCCAAGACTCGGGTGGTCC
AGGAGCACGGCGGGCGGGCGGTGATCATCTCTGACAACGCAGTTGACAATGACAGCTTCTAC
GTGGAGATGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCGCCCTCTTCCTGCTCGG
CCGAGACGGCTACATGATCCGCCGCTCTCTGGAACAGCATGGGCTGCCATGGGCCATCATTT
CCATCCCAGTCAATGTCACCAGCATCCCCACCTTTGAGCTGCTGCAACCGCCCTGGACCTTC
TGG**TAGA**AAGAGTTTGTCCCACATTCCAGCCATAAGTGACTCTGAGCTGGGAAGGGGAAACCC
AGGAATTTTGCTACTTGGAATTTGGAGATAGCATCTGGGGACAAGTGGAGCCAGGTAGAGGA
AAAGGGTTTGGGCGTTGCTAGGCTGAAAGGGAAGCCACACCACTGGCCTTCCCTTCCCCAGG
GCCCCAAGGGTGTCTCATGCTACAAGAAGAGGCAAGAGACAGGCCCCAGGGCTTCTGGCTA
GAACCCGAAACAAAAGGAGCTGAAGGCAGGTGGCCTGAGAGCCATCTGTGACCTGTCACACT
CACCTGGCTCCAGCCTCCCCTACCCAGGGTCTCTGCACAGTGACCTTCACAGCAGTTGTTGG
AGTGGTTTAAAGAGCTGGTGTGTTGGGGACTCAATAAACCCCTCACTGACTTTTTTAGCAATAAA
GCTTCTCATCAGGGTTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 220

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76532

><subunit 1 of 1, 188 aa, 1 stop

><MW: 21042, pI: 5.36, NX(S/T): 2

MVPGAAGWCCLVLWLPACVAAHGFRIDYLYFQVLSPGDIRYIFTATPAKDFGGIFHTRYEQ
IHLVPAEPPEACGELSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVIISDNAVDNDSF
YVEMIQDSTQRTADIPALFLLGRDGYMIRRSLEQHGLPWAIISIPVNVTSIPTFELLQPPWTFW

Signal peptide:

amino acids 1-20

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 221

TCTGCCTCCACTGCTCTGTGCTGGGATCATGGAACTTGCACTGCTGTGTGGGCTGGTGGTGA
TGGCTGGTGTGATTCCAATCCAGGGCGGGATCCTGAACCTGAACAAGATGGTCAAGCAAGTG
ACTGGGAAAATGCCCATCCTCTCCTACTGGCCCTACGGCTGTCCTGCGGACTAGGTGGCAG
AGGCCAACCCAAAGATGCCACGGACTGGTGCTGCCAGACCCATGACTGCTGCTATGACCAC
TGAAGACCCAGGGGTGCGGCATCTACAAGGACAACAACAAAAGCAGCATACATTGTATGGAT
TTATCTCAACGCTATTGTTTAATGGCTGTGTTTAATGTGATCTATCTGGAAAATGAGGACTC
CGAATAAAAAGCTATTACTAWTTNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 222

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76538
><subunit 1 of 1, 116 aa, 1 stop
><MW: 12910, pI: 6.41, NX(S/T): 1
MELALLCGLVVMAGVPIQGGILNLNKMVKQVTGKMPILSYWPYGCHCGLGGRGQPKDATDW
CCQTHDCCYDHLKTQGCGIYKDNNKSSIHCMDSLQRYCLMAVFNVIIYLENEDSE
```

Important features of the protein:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 1-24

N-glycosylation site.

amino acids 86-89

N-myristoylation sites.

amino acids 20-25, 45-50

Phospholipase A2 histidine active site.

amino acids 63-70

FIGURE 223

CTCGCTTCTTCCTTCTGGATGGGGGGCCAGGGGGCCAGGAGAGTATAAAGGCGATGTGGAG
GGTGCCCGGCACAACCAGACGCCAGTCACAGGCGAGAGCCCTGGG**ATG**CACCGGCCAGAGG
CCATGCTGCTGCTGCTCACGCTTGCCCTCCTGGGGGGCCCCACCTGGGCAGGGAAGATGTAT
GGCCCTGGAGGAGGCAAGTATTTACAGCACCCTGAAGACTACGACCATGAAATCACAGGGCT
GCGGGTGTCTGTAGGTCTTCTCCTGGTGAAAAGTGTCCAGGTGAAACTTGGAGACTCCTGGG
ACGTGAAACTGGGAGCCTTAGGTGGGAATACCCAGGAAGTCACCCTGCAGCCAGGCGAATAC
ATCACAAAAGTCTTTGTGCGCTTCCAAGCTTTCCTCCGGGGTATGGTCATGTACACCAGCAA
GGACCGCTATTTCTATTTTGGGAAGCTTGATGGCCAGATCTCCTCTGCCTACCCCAGCCAAG
AGGGGCAGGTGCTGGTGGGCATCTATGGCCAGTATCAACTCCTTGGCATCAAGAGCATTGGC
TTTGAATGGAATTATCCACTAGAGGAGCCGACCACTGAGCCACCAGTTAATCTCACATACTC
AGCAAACCTCACCCGTGGGTCGC**TAG**GGTGGGGTATGGGGCCATCCGAGCTGAGGCCATCTGT
GTGGTGGTGGCTGATGGTACTGGAGTAACTGAGTCGGGACGCTGAATCTGAATCCACCAATA
AATAAAGCTTCTGCAGAAAA

FIGURE 224

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76541
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19600, pI: 5.89, NX(S/T): 1
MHRPEAMLLLLLTLALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSVGLLLVKSQVVK
LGDSWDVKLGALGGNTQEVTLPGEYITKVVFVAFQAFLRGMVMYTSKDRYFYFGKLDGQISS
AYPSQEGQVLVGIYGQYQLLGKISIGFEWNYPLEEPTTEPPVNLTYSANSPVGR
```

Signal peptide:

amino acids 1-22

GCTTGAGCGTGTGCGCGGGTACGCGGGCTCTCTCTGCCTTCTGCGGCTCCAACGCAGCTCTGTGGCT
GAACTGGGTGCTCATCACGGGAACTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAG
CCCCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTT
TTTTTTTTTAACCGCCCCCTCCCCACCCCCCAAAAAAAGCTGTAAAGATGCAAAAACGTAATAT
CCATGAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATT
TATTTGTTCTTGGAGTGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCT
CCCAAGGGGTCCAATTTTTCTTCTGCGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTC
ACAGGGGCTGTGCATGCAACTGGCCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAA
CAATACAAAGGATGGGTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTT
ATAGCCCCCACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTG
TAGGTGTGAAGGCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCTCAAGTA
TATCTGCTGTTGCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAACCTAAGTATAAT
CAATTTAAAGGGCTCAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATAT
TGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCCAATAGAA
TCTCCTATTTTCTTAACAATAACCTTCAGACCTGTGACAAATTTACGGAACCTGGATCTGTCC
TATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTGCGGAAGCTGCTGAGTTT
ACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACC
TGGAACTTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGC
ATGATCAGACTCAAAGAAGCTTCACCTGGAGCACAAATCAATTTTTCCAAGCTCAACCTGGCCCT
TTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGTGCATAG
GACAGACCATGTCTTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATC
GAAGCTTTTCAGTGGACCCAGTGTTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGA
TTCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACA
TCAGTCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAGCTGGCTG
AAAAGTTTTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGG
AGTAAATGTGATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGT
TTGATCTGGCCAGGGCTCTCCCAAAGCCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCAT
GAGAGCAAACCCCTTTGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGGATGC
TGACGCCGAGCACATCTCTTCCATAAAATCATCGCGGGCAGCGTGCGGCTTTTCTGTCCG
TGCTCGTCATCTGCTGTTATCTACGTGTCAATGGAAGCGGTACCCTGCGAGCATGAAGCAG
CTGCAGCAGCGCTCCCTCATGCGAAGGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAAT
GACTCCCAGCACCCAGGAATTTTATGTAGATTATAAACCACCAACACGGAGACCAGCGAGA
TGCTGCTGAATGGGACGGGACCCTGCACCTATAACAAATCGGGCTCCAGGGAGTGTGAGGTA
TGAACCATTGTGATAAAAAGAGCTCTTAAAAGCTGGGAAATAAGTGGTGCTTTATTGAACTC
TGGTGACTATCAAGGGAACGCGATGCCCCCCCTCCCTTCCCTCTCCCTCTCACTTTGGTG
CAAGATCCTTCCTTGTCCGTTTTAGTGCATTCATAATACTGGTCATTTTCCCTCTCATACATA
ATCAACCCATTGAAAATTTAAATACCACAATCAATGTGAAGCTTGAAGCTCCGGTTTTAATATAA
TACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTCGCATTTGTTTTAAGATAAACT
TCTTTCATAGGTAAAAA

FIGURE 226

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77301
><subunit 1 of 1, 513 aa, 1 stop
><MW: 58266, pI: 9.84, NX(S/T): 4
MGFNVIRLLSGSAVALVIAPTPLLTLMLSSAERGCPKGCRCCEGKMVYCESQKLQEIPSSISAG
CLGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYF
LNNTFRPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELL
DLGYNRIRSLARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQONLYLQWNKISVIGQTM
SWTWSSLQRLDLSGNEIEAFSGPSVFQCVPNLQRLNLDLSDNKLTFIGQEILDSWISLNDISLA
GNIWECSRNICSLVNWLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLA
RALPKPTFKPKLPRPKHESKPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVI
LLVIYVSWKRYPASMKQLQQRSLMRRHRKKKRQSLKQMPSTQEFYVDYKPTNTTETSEMLLN
GTGPCTYNKSGSRECEV

Important features of the protein:**Signal peptide:**

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 227

AGTTCTGAGAAAGAAGGAAATAAACACAGGCACCAAACCACTATCCTAAGTTGACTGTCCTT
 TAAAT**ATGT**CAAGATCCAGACTTTTCAGTGTACCTCAGCGATCTCAACGATAGGGATCTTG
 TGTTTGCCGCTATTCCAGTTGGTGCTCTCGGACCTACCATGCGAAGAAGATGAAATGTGTGT
 AAATTATAATGACCAACACCCCTAATGGCTGGTATATCTGGATCCTCCTGCTGCTGGTTTTGG
 TGGCAGCTCTTCTCTGTGGAGCTGTGGTCCTCTGCCTCCAGTGCTGGCTGAGGAGACCCCGA
 ATTGATTCTCACAGGCGCACCATGGCAGTTTTTGTCTGTTGGAGACTTGGACTCTATTTATGG
 GACAGAAGCAGCTGTGAGTCCAACTGTTGGAATTCACCTTCAAACCTCAAACCCCTGACCTAT
 ATCCTGTTCCCTGCTCCATGTTTTGGCCCTTTAGGCTCCCCACCTCCATATGAAGAAATTGTA
 AAAACAACCT**TGA**TTTTAGGTGTGGATTATCAATTTAAAGTATTAACGACATCTGTAATTCCA
 AAACATCAAATTTAGGAATAGTTATTTTCAGTTGTTGGAAATGTCCAGAGATCTATTCATATA
 GTCTGAGGAAGGACAATTTCGACAAAAGAATGGATGTTGGAAAAAATTTTGGTCATGGAGATG
 TTAAATAGTAAAGTAGCAGGCTTTTGATGTGTCACTGCTGTATCATACTTTTATGCTACAC
 AACCAAATTAATGCTTCTCCACTAGTATCCAAACAGGCAACAATTAGGTGCTGGAAGTAGTT
 TCCATCACATTTAGGACTCCACTGCAGTATACAGCACACCATTTTCTGCTTTAAACTCTTTC
 CTAGCATGGGGTCCATAAAAATTATTATAATTTAACAATAGCCCAAGCCGAGAATCCAACAT
 GTCCAGAACCAGAACCAGAAAGATAGTATTTGAATGAAGGTGAGGGGAGAGAGTAGGAAAAA
 GAAAAGTTTGGAGTTGAAGGGTAAAGGATAAATGAAGAGGAAAAGGAAAAGATTACAAGTCT
 CAGCAAAAACAAGAGGTTTTATGCCCAACCTGAAGAGGAAGAAATTGTAGATAGAAGGTGA
 AGGAGATTGCTGAAGATATAGAGCACATATAATGCCAACACGGGGAGAAAAGAAAATTTCCC
 CTTTTACAGTAATGAATGTGGCCTCCATAGTCCATAGTGTTTCTCTGGAGCCTCAGGGCTTG
 GCATTTATTGCAGCATCATGCTAAGAACCTTCGGCATAGGTATCTGTTCCCATGAGGACTGC
 AGAAGTAGCAATGAGACATCTTCAAGTGGCATTTTTGGCAGTGGCCATCAGCAGGGGGACAGA
 CAAAACATCCATCACAGATGACATATGATCTTCAGCTGACAAATTTGTTGAACAAAACAAT
 AAACATCAATAGATATCTAAAAA

FIGURE 228

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77303

><subunit 1 of 1, 146 aa, 1 stop

><MW: 16116, pI: 4.99, NX(S/T): 0

MSRSRLFSVTSIAISTIGILCLPLFQLVLSDLPCCEEDEMCVNYNDQHPNGWYIWILLLLVLVA

ALLCGAVVLCLQCWLRRPRIDSHRRTMAVFAVGDLDSIYGTEAAVSPTVGIHLQTQTPDLYP

VPAPCFGPLGSPPPYEEIVKTT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 52-70

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 229

GAGCGGAGTAAAATCTCCACAAGCTGGGAACAAACCTCGTCCCAACTCCCACCCACCGGCGT
 TTCTCCAGCTCGATCTGGAGGCTGCTTCGCCAGTGTGGGACGCAGCTGACGCCCCGCTTATTA
 GCTCTCGCTGCGTCGCCCCGGCTCAGAAGCTCCGTGGCGGCGGCGACCGTGACGAGAAGCCC
 ACGGCCAGCTCAGTTCTCTTCTACTTTGGGAGAGAGAGAAAAGTCAGATGCCCCCTTTAAACT
 CCCTCTTCAAACTCATCTCCTGGGTGACTGAGTTAATAGAGTGGATACAACCTTGCTGAAG
 ATGAAGAATATACAATATTGAGGATATTTTTTTCTTTTTTTTTTCAAGTCTTGATTTGTGGC
 TTACCTCAAGTTACCATTTTTTCAGTCAAGTCTGTTTGTTTGCTTCTTCAGAAATGTTTTTA
 CAATCTCAAGAAAAAATATGTCCCAGAAATTGAGTTTACTGTTGCTTGTATTTGGACTCAT
 TGGGGATTGATGTTACTGCACTATACTTTTCAACAACCAAGACATCAAAGCAGTGTCAAGTT
 ACGTGAGCAAATACTAGACTTAAGCAAAAGATATGTTAAAGCTCTAGCAGAGGAAAATAAGA
 ACACAGTGGATGTGCGAAGCGGTGCTTCTATGGCAGGATATGCGGATCTGAAAAGAACAATT
 GCTGTCCTTCTGGATGACATTTTGCAACGATTGGTGAAGCTGGAGAACAAAGTTGACTATAT
 TGTTGTGAATGGCTCAGCAGCCAACACCACCAATGGTACTAGTGGGAATTTGGTGCCAGTAA
 CCACAAATAAAAGAACGAATGTCTCGGGCAGTATCAGATAGCAGTTGAAAATCACCTTGTGC
 TGCTCCATCCACTGTGGATTATATCCTATGGCAGAAAAGCTTTATAATTGCTGGCTTAGGAC
 AGAGCAATACTTTACAATAAAAGCTCTACACATTTTCAAGGAGTATGCTGGATTCATGGAAC
 TCTAATTCTGTACATAAAAAATTTTAAAGTTATTTGTTTGCTTTCAGGCAAGTCTGTTCAATG
 CTGTACTATGTCCTTAAAGAGAATTTGGTAACTTGGTTGATGTGGTAAGCAGATAGGTGAGT
 TTTGTATAAATCTTTTGTGTTTGAGATCAAGCTGAAATGAAAACACTGAAAAACATGGATT
 ATTTCTATAACACATTTATTTAAGTATATAACACGTTTTTTTGGACAAGTGAAGAATGTTTAA
 TCATTCTGTCAATTTGTTCTCAATAGATGTAAGTGTAGACTACGGCTATTTGAAAAATGTG
 CTTATTGTACTATATTTTGTATTCCAATTATGAGCAGAGAAAGGAAATATAATGTTGAAAA
 TAATGTTTTGAAATCATGACCCAAAGAATGTATTGATTTGCACTATCCTTCAGAATAACTGA
 AGGTTAATTATTGTATATTTTTTAAAAATTACACTTATAAGAGTATAATCTTGAAATGGGTAG
 CAGCCACTGTCCATTACCTATCGTAAACATTTGGGGCAATTTAATAACAGCATTTAAATAGTT
 GTAAACTCTAATCTTATACTTATTGAAGAATAAAAGATATTTTTTATGATGAGAGTAACAATA
 AAGTATTCATGATTTTTTCACATACATGAATGTTTCAATTTAAAAGTTTAAATCCTTTGAGTGTCT
 ATGCTATCAGGAAAGCACATTATTTCCATATTTGGGTAAATTTTGCTTTTATTATATTGGTC
 TAGGAGGAAGGGACTTTGGAGAATGGAACCTTGAGGACTTTAGCCAGGTGTATATAATAAA
 GGTACTTTTTGTGCTGCATTAAATTGCTTGGAAGTGTTAACATTATATTATATAAGAGTATC
 CTTTATGAAATTTTGAATTTGTATAACAGATGCATTAGATATTCATTTTATATAATGGCCAC
 TTAATAAAGAACATTTAAATATAAACTATGAAGATTGACTATCTTTTCAGGAAAAAAGCT
 GTATATAGCACAGGGAACCCTAATCTTGGGTAAATCTAGTATAAAACAAATTATACTTTTAT
 TTAAATTTCCCTTGTAGCAAATCTAATTGCCACATGGTGCCCTATATTTTCATAGTATTTAT
 CTCTATAGTAACTGCTTAAAGTGCAGCTAGCTTCTAGATTTAGACTATATAGAATTTAGATAT
 TGTATTGTTTCGTCATTATAATATGCTACCACATGTAGCAATAATTACAATATTTTATTAATA
 TAAATATGTGAAATATTGTTTCATGAAAGACAGATTTCCAAATCTCTCTCTCTCTCTGTA
 CTGTCTACCTTTATGTGAAGAAATTAATTATATGCCATTGCCAGGT

FIGURE 230

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77648

><subunit 1 of 1, 140 aa, 1 stop

><MW: 15668, pI: 10.14, NX(S/T): 5

MFFTISRKNMSQKLSLLLLVFGLIWGLMLLHYTFQQPRHQSSVKLREQILDLSKRYVKALAE
ENKNTVDVENGASMAGYADLKRTIAVLLDDILQRLVKLENKVDYIVVNGSAANTTNGTSGNL
VPVTTNKRTNVSGSIR

Important features of the protein:

Signal peptide:

amino acids 1-26

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 231

CGCGGCCGGGCCGCCGGGGTGAAGCGTGCCGAGGCGGCTGTGGCGCAGGCTTCCAGCCCCAC
CATGCCGTGGCCCCCTGCTGCTGCTGCTGGCCGTGAGTGGGGCCCAGACAACCCGGCCATGCT
TCCCCGGGTGCCAATGCGAGGTGGAGACCTTCGGCCTTTTCGACAGCTTCAGCCTGACTCGG
GTGGATTGTAGCGGCCTGGGCCCCCACATCATGCCGGTGCCCATCCCTCTGGACACAGCCCA
CTTGACCTGTCCTCCAACCGGCTGGAGATGGTGAATGAGTCGGTGTTGGCGGGGCCGGGCT
ACACGACGTTGGCTGGCCTGGATCTCAGCCACAACCTGCTCACCAGCATCTCACCCACTGCC
TTCTCCCGCCTTCGCTACCTGGAGTCGCTTGACCTCAGCCACAATGGCCTGACAGCCCTGCC
AGCCGAGAGCTTCACCAGCTCACCCTGAGCGACGTGAACCTTAGCCACAACCAGCTCCGGG
AGGTCTCAGTGTCTGCCTTCACGACGCACAGTCAGGGCCGGGCACTACACGTGGACCTCTCC
CACAACCTCATTCACCGCCTCGTGCCCCACCCACGAGGGCCGGCCTGCCTGCGCCACCAT
TCAGAGCCTGAACCTGGCCTGGAACCGGCTCCATGCCGTGCCAACCTCCGAGACTTGCCCC
TGCGCTACCTGAGCCTGGATGGGAACCTCTAGCTGTCAATTGGTCCGGGTGCCTTCGCGGGG
CTGGGAGGCCTTACACACCTGTCTCTGGCCAGCCTGCAGAGGCTCCCTGAGCTGGCGCCAG
TGGCTTCCGTGAGCTACCGGGCCTGCAGGTCTTGACCTGTGCGGGCAACCCCAAGCTTAACT
GGGCAGGAGCTGAGGTGTTTTAGGCCTGAGCTCCCTGCAGGAGCTGGACCTTTCGGGCACC
AACCTGGTGCCCCCTGCCTGAGGCGCTGCTCCTCCACCTCCCGGCACTGCAGAGCGTCAGCGT
GGGCCAGGATGTGCGGTGCCGGCGCCTGGTGCGGGAGGGCACCTACCCCGGAGGCCTGGCT
CCAGCCCCAAGGTGCCCTGCACTGCGTAGACACCCGGGAATCTGCTGCCAGGGGCCCCACC
ATCTT**GTGA**CAAATGGTGTGGCCCAGGGCCACATAACAGACTGCTGTCTCTGGGCTGCCTCAG
GTCCCGAGTAACCTTATGTTCAATGTGCCAACACCAAGTGGGGAGCCCGCAGGCCTATGTGGCA
GCGTCACCACAGGAGTTGTGGGCCTAGGAGAGGCTTTGGACCTGGGAGCCACACCTAGGAGC
AAAGTCTCACCCCTTTGTCTACGTTGCTTCCCCAAACCATGAGCAGAGGGACTTCGATGCCA
AACCAGACTCGGGTCCCCTCCTGCTTCCCTTCCCCACTTATCCCCCAAGTGCCTTCCCTCAT
GCCTGGGCGGCCTGACCCGCAATGGGCAGAGGGTGGGTGGGACCCCTGCTGCAGGGCAGA
GTTTCAGGTCCACTGGGCTGAGTGTCCCCTTGGGCCCATGGCCAGTCACTCAGGGGCGAGTT
TCTTTTCTAACATAGCCCTTTCTTTGCCATGAGGCCATGAGGCCCGCTTCATCCTTTTCTAT
TTCCCTAGAACCTTAATGGTAGAAGGAATTGCAAAGAATCAAGTCCACCCTTCTCATGTGAC
AGATGGGGAAACTGAGGCCTTGAGAAGGAAAAAGGCTAATCTAAGTTCCCTGCGGGCAGTGGC
ATGACTGGAGCACAGCCTCCTGCCTCCAGCCCGGACCCAATGCACTTTCTTGTCTCCTCTA
ATAAGCCCCACCCTCCCCGCCTGGGCTCCCCTTGCTGCCCTTGCTGTTCCTTATGACACA
GGAGTAGCAGCAGCAGGACAGGCAAGAGCCTCACAAGTGGGACTCTGGGCCTCTGACCAGCT
GTGCGGCATGGGCTAAGTCACTCTGCCCTTCGGAGCCTCTGGAAGCTTAGGGCACATTGGTT
CCAGCCTAGCCAGTTTCTCACCCCTGGGTGGGGTCCCCCAGCATCCAGACTGGAAACCTACC
CATTTTCCCCTGAGCATCCTCTAGATGCTGCCCAAGGAGTTGCTGCAGTTCTGGAGCCTCA
TCTGGCTGGGATCTCCAAGGGGCCTCCTGGATTCACTCCCCACTGGCCCTGAGCACGACAGC
CCTTCTTACCCTCCCAGGAATGCCGTGAAAGGAGACAAGGTCTGCCCGACCCATGTCTATGC
TCTACCCCCAGGGCAGCATCTCAGCTTCCGAACCTGGGCTGTTTCCTTAGTCTTCATTTTA
TAAAAGTTGTTGCCTTTTTTAACGGAGTGTCACTTTCAACCGGCCTCCCCTACCCCTGCTGGC
CGGGGATGGAGACATGTCATTTGTAAAAGCAGAAAAAGGTTGCATTTGTTCACCTTTGTAAAT
ATTGTCCTGGGCCTGTGTTGGGGTGTGGGGGAAGCTGGGCATCAGTGGCCACATGGGCATC
AGGGGCTGGCCCCACAGAGACCCACAGGGCAGTGAGCTCTGTCTTCCCCACCTGCCTAGC
CCATCATCTATCTAACCGGTCCTTGATTTAATAAACACTATAAAAGGTTTAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77652

><subunit 1 of 1, 353 aa, 1 stop

><MW: 37847, pI: 6.80, NX(S/T): 2

MPWPLLLLLLAVSGAQTRPCFPGCQCEVETFGLFDSFSLTRVDCSGLGPHIMPVPIPLDTAH
LDLSSNRLEMVNESVLAGPGYTTLAGLDLSHNLLTSISPTAFSRLRYLESLDLSHNGLTALP
AESFTSSPLSDVNLSHNQLREVSVSAFTTHSQGRALHVDLSHNLIHRLVPHPTRAGLPAPTI
QSLNLAWNRLHAVPNLRDLPLRYLSLDGNPLAVIGPGAFAFAGLGGLTHLSLASLQRLPELAPS
GFRELPGQLQVLDLSGNPKLNWAGAEVFSGLSSQLQLDLSGTNLVPLPEALLLHLPALQSVSV
GQDVRCRRLVREGTYPRRPGSSPKVPLHCVDTRESAARGPTIL

Signal peptide:

amino acids 1-16

Transmembrane domains:

amino acids 215-232, 287-304

FIGURE 233

GATGGCGCAGCCACAGCTTCTGTGAGATTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGG
 GGACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTAT**ATG**CGTCAATCCCCA
 AAACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCT
 GTTCCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGC
 CACGACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCAC
 GGGAGGCTTGGCAGTTTTTTCTTACTCCTGTGGTCTCCAGATTTCAAGGCCTAAGATGAAAGCC
 TCTAGTCTTGCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGG
 ACTGAAGACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATG
 GATTTTCTGAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTA
 AGGAGGACTGAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTT
 GCTAAGACTCTATCTGGACAGGGTATTTAAAAACTACCAGACCCCTGACCATTATACTCTCC
 GGAAGATCAGCAGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCAT
 GCCCACATGACATGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCA
 CTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAACTAGACATTCTTC
 TGCAATGGATGGAGGAGACAGAA**TAG**GAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGT
 CAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACT
 GTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGAT
 TGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTTTTGTAATATCTT
 TCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTT
 ATTTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTGACTAG
 AGCAGGTGATGTATTTTTTATACAGTAAAAAATAACCTTGTAATTCTAGAAGAGTGGCT
 AGGGGGGTTATTCAATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGA
 TATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTTGATGTGGAATT
 GCACATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTAT
 CTTCCAGCCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATA
 CCAAAAAAAAAAAAAAAAAAAAA

FIGURE 234

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83500
><subunit 1 of 1, 261 aa, 1 stop
><MW: 29667, pI: 8.76, NX(S/T): 0
MRQFPKTSFDISPEMSFSIYSLQVPAVPGGLTCWALTAEPGWGQNKGATTCATNSHSDSELRP
EIFSSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNL
QEIRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTP
DHYTLRKISSLANSLFTIKKDLRLSHAHMTCHCGEEAMKKSSQILSHFEKLEPQAAVVKALG
ELDILLQWMEETE
```

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 235

CCGTTATCGTCTTGCGCTACTGCTGA**ATG**TCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCG
CTGACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGC
CGAGCTAGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAG
CTCTTGCTCGGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACA
GCCCTAGGGATCATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCAT
TTACAGACACGTAGTGTATTCTGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTG
TGTTTGGCAAAAGTGAAGATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATG
GCTGGTGTTATTGGCCAGTTTTTTAGCCAATCCAAGTACCTAGTGAAGGTTTCAAGATGCAAAAT
GGAAGGAAAAAGGAAACTGGAAGGAAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTG
CAAAAATCTTAGCTGAAGGAGGAATACGAGGGCTTTGGGCAGGCTGGGTACCCAATATACAA
AGAGCAGCACTGGTGAATATGGGAGATTTAACCACCTTATGATACAGTGAAACACTACTTGGT
ATTGAATACACCACTTGAGGACAATATCATGACTCACGGTTTATCAAGTTTATGTTCTGGAC
TGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCATCAAAGCAGAATAATGAATCAACCA
CGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCGACTGACTGCTTGATTGAGGCTGT
TCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTTTACCATCTTGGCTGAGAATGACCC
CTTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGATGAGTGGAGTCAGTCCA
TTTT**TAA**

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77568

><subunit 1 of 1, 323 aa, 1 stop

><MW: 36064, pI: 9.33, NX(S/T): 1

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEEAALARLGDGA
RESAPYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEH
YPLWKSVIGGMMAGVIGQFLANPTDLVKVQMMEGKRRKLEGKPLRFRGVHHAFAKILAEGGI
RGLWAGWVPNIQRAALVNMGDLTTYDTVKHYLVNLNTPLEDNIMTHGLSSLCSGLVASILGTP
ADVIKSRIMNQPRDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLT
YEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 237

CGGACGCGTGGGCGCGGGACGCCGGCAGGGTTGTGGCGCAGCAGTCTCCTTCCTGCGCGCGC
GCCTGAAGTCGGCGTGGGCGTTTGAGGAAGCTGGGATACAGCATTTAATGAAAAATTTATGC
TTAAGAAGTAAAA**ATG**GCAGGCTTCCTAGATAATTTTCGTTGGCCAGAATGTGAATGTATTG
ACTGGAGTGAGAGAAGAAATGCTGTGGCATCTGTTGTCGCAGGTATATTGTTTTTTACAGGC
TGGTGGATAATGATTGATGCAGCTGTGGTGTATCCTAAGCCAGAACAGTTGAACCATGCCTT
TCACACATGTGGTGTATTTCCACATTGGCTTTCTTCATGATAAATGCTGTATCCAATGCTC
AGGTGAGAGGTGATAGCTATGAAAGCGGCTGTTTAGGAAGAACAGGTGCTCGAGTTTGGCTT
TTCATTGGTTTTCATGTTGATGTTTGGGTCACTTATTGCTTCCATGTGGATTCTTTTTTGGTGC
ATATGTTACCCAAAATACTGATGTTTATCCGGGACTAGCTGTGTTTTTTCAAATGCACTTA
TATTTTTTTAGCACTCTGATCTACAAATTTGGAAGAACCGAAGAGCTATGGAC**TC**GATCAC
TTCTTAAGTCACATTTTCCTTTTGTATATTCTGTTTGTAGATAGGTTTTTTATCTCTCAGT
ACACATTGCCAAATGGAGTAGATTGTACATTAAATGTTTTGTTTCTTTACATTTTTTATGTTT
TGAGTTTTGAAATAGTTTTATGAAATTTCTTTATTTTTTCATTGCATAGACTGTTAATATGTA
TATAATACAAGACTATATGAATTGGATAATGAGTATCAGTTTTTTATTCTTGAGATTTAGAA
CTTGATCTACTCCCTGAGCCAGGGTTACATCATCTTGTCATTTTAGAAAGTAACCACTCTTGT
CTCTCTGGCTGGGCACGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGG
CCGATTGCTTGAGGTCAAGTGTGAGACCAGCCTGGCCAACATGGCGAAACCCCATCTACT
AAAAATACAAAAATTAGCCAGGCATGGTGGTGGGTGCCTGTAATCCCAGCTACCTGGGAGGC
TGAGGCAGGAGAATCGCTTGAACCCGGGGGGCAGAGGTTGCAGTGAGCTGAGTTTGCGCCAC
TGCACTCTAGCCTGGGGGAGAAAGTGAACTCCCTCTCAAAAAAAGACCACTCTCAGTATC
TCTGATTTCTGAAGATGTACAAAAAATATAGCTTCATATATCTGGAATGAGCACTGAGCCA
TAAAAGGTTTTCAGCAAGTTGTAACCTATTTTGGCCTAAAAATGAGGTTTTTTTGGTAAAGA
AAAAATATTTGTTCTTATGTATTGAAGAAGTGACTTTTATATAATGATTTTTTAAATGCCC
AAAGGACTAGTTTGAAAGCTTCTTTTAAAAAGAATTCCTCTAATATGACTTTATGTGAGAA

FIGURE 238

MAGFLDNFRWPECECIDWSERRNAVASVVGILFFTGWIMIDAADVYPKPEQLNHAFHTCG
VFSTLAFFMINAVSNAQVRGDSYESGCLGRTGARVWLFIGFMLMFGSLIASMWILEFGAYVTQ
NTDVYPGLAVFFQNALIFFSTLIYKFG RTEELWT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 23-42 (type II), 60-80, 97-117, 128-148

Downloaded from www.sagepub.com at [university] on [date]

FIGURE 239

GTTGATGGCAAACCTTCCTCAAAGGAGGGGCAGAGCCTGCGCAGGGCAGGAGCAGCTGGCCCA
CTGGCGGGCCCGCAACACTCCGTCTCACCCCTCTGGGCCCCACTGCATCTAGAGGAGGGCCGTCT
GTGAGGCCACTACCCCTCCAGCAACTGGGAGGTGGGACTGTCAGAAGCTGGCCCAGGGTGGT
GGTCAGCTGGGTGAGGGACCTACGGCACCTGCTGGACCACCTCGCCTTCTCCATCGAAGCAG
GGAAGTGGGAGCCTCGAGCCCTCGGGTGGAAGCTGACCCCAAGCCACCCTTCACCTGGACAG
GATGGAGAGTGTGAGGTGTGCTTCGCCTCCTGGCCCTCATCTTTGCCATAGTCACGACATGGA
TGTTTATTCTGAAGCTACATGAGCTTCAGCATGAAAACCATCCGTCTGCCACGCTGGCTGGCA
GCCTCGCCCCACCAAGGAGATCCAGGTAAAGTACAAGTGTGGCCTCATCAAGCCCTGCCC
AGCCAACTACTTTGCGTTTAAATCTGCAGTGGGGCCGCCAACGTCGTGGGCCCTACTATGT
GCTTTGAAGACCGCATGATCATGAGTCCTGTGAAAAACAATGTGGGCAGAGGCCTAAACATC
GCCCTGGTGAATGGAACCACGGGAGCTGTGCTGGGACAGAAGGCATTTGACATGTACTCTGG
AGATGTTATGCACCTAGTGAAATTCCTTAAAGAAATTCGGGGGGTGCAGTGGTGTGGTGG
CCTCCTACGACGATCCAGGGACCAAATGAACGATGAAAGCAGGAAACTCTTCTCTGACTTG
GGGAGTTCCTACGCAAAACAACTGGGCTTCCGGGACAGCTGGGTCTTCATAGGAGCCAAAGA
CCTCAGGGGTAAAAGCCCCCTTTGAGCAGTTCTTAAAGAACAGCCAGACACAAACAAATACG
AGGGATGGCCAGAGCTGCTGGAGATGGAGGGCTGCATGCCCCGAAGCCATTT**TAG**GGTGGC
TGTGGCTCTTCCTCAGCCAGGGGCCTGAAGAAGCTCCTGCCTGACTTAGGAGTCAGAGCCCC
GCAGGGGCTGAGGAGGAGGAGCAGGGGGTGTGCGTGGAAGGTGCTGCAGGTCCTTGACGC
TGTGTGCGCCTCTCCTCCTCGGAAACAGAACCTCCCACAGCACATCCTACCCGGAAGACC
AGCCTCAGAGGGTCCTTCTGGAACCAGCTGTCTGTGGAGAGAATGGGGTGCTTTCGTCAGGG
ACTGCTGACGGCTGGTCCTGAGGAAGGACAACTGCCAGACTTGAGCCCAATTAAATTTTA
TTTTTGCTGGTTTTGAAAAAAAAAAAAAAAAAAAAA



FIGURE 240

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59814

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24963, pI: 9.64, NX(S/T): 1

MRVSGVLRLLALIFAIVTTWMFIRSYMFSMKTIRLPRWLAASPTKEIQVKKYKCGLIKPCP
ANYFAFKICSGAANVVGPTMCFEDRMIMSPVKNNVGRGLNIALVNGTTGAVLGQKAQDMYSG
DVMHLVKFLKEIPGGALVLVASYYDDPGTKMNDESRLKLFSDLGSSYAKQLGFRDSWVFIGAKD
LRGKSPFEQFLKNSPDTNKEYEGWPELLEMEGCMPPKPF

Important features:**Signal peptide:**

amino acids 1-15

ATP/GTP-binding site motif A (P-loop).

amino acids 184-191

N-glycosylation site.

amino acids 107-110

FIGURE 241

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGAT
 CCAGAAACCCATGATACCCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACA
 GCAAGAGAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCTC
 TCACTCCTCCCTCCCTCTCTCTCTGCTGTCCTAGTCCTCTAGTCCTCAAATTTCCAGTCCC
 CTGCACCCCTTCCTGGGACACTATGTTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCC
 TGGCTGCAGATGGGGGTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTTGGCCA
 GCCTCTTACCCTGAGTGTGGAAACAATGCCCAGTCGCCCATCGATATTCAGACAGACAGTGT
 GACATTTGACCCTGATTTGCCTGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGC
 CTTTGGACCTGCACAACAATGGCCACACAGTGCAACTCTCTCTGCCCTCTACCCTGTATCTG
 GGTGGACTTCCCCGAAAATATGTAGCTGCCCAGCTCCACCTGCACTGGGGTCAGAAAGGATC
 CCCAGGGGGGTGAGAACACCAGATCAACAGTGAAGCCACATTTGCAGAGCTCCACATTGTAC
 ATTATGACTCTGATTCCTATGACAGCTTGAGTGAGGCTGCTGAGAGGCCTCAGGGCCTGGCT
 GTCTTGGGCATCCTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTCTGAG
 TCACTTGCAATGAAGTCAGGCATAAAGATCAGAAGACCTCAGTGCCTCCCTTCAACCTAAGAG
 AGCTGCTCCCCAAACAGCTGGGGCAGTACTTCCGCTACAATGGCTCGCTCACAACCTCCCCCT
 TGCTACCAGAGTGTGCTCTGGACAGTTTTTTTATAGAAGGTCCCAGATTTCAATGGAACAGCT
 GGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAAGCTTCTGGTACAGA
 ACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTTCATCCAAGCAGGA
 TCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGCTGTCTCTG
 CCTTCTCCTGGCTGTTTTATTTTCATTGCTAGAAAGATTTCGGAAGAAGAGGCTGGAAAACCGAA
 AGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCATAAATTCCTTCTCAGATACCA
 TGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTG
 GCCAGAAACACTGTAGGAGTAGTAAGCAGATGTCTCCTTCCCCTGGACATCTCTTAGAGAG
 GAATGGACCCAGGCTGTCATTCCAGGAAGAAGTGCAGAGCCTTCAGCCTCTCCAAACATGTA
 GGAGGAAATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAACCTCTGTTTAGTTGCAGGG
 GAAGTTTGGGATATACCCCAAAGTCTCTACCCCTCACTTTTATGGCCCTTTCCCTAGATA
 TACTGCGGGATCTCTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATA
 TATTTGGAAATTAAAGTTTCTGACTTT

FIGURE 242

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62812

><subunit 1 of 1, 337 aa, 1 stop

><MW: 37668, pI: 6.27, NX(S/T): 1

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWHPASYPECGNNAQSPIDIQTDSVTFDPDLP
ALQPHGYDQPGTEPLDLHNNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQ
INSEATFAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRH
KDQKTSVPPFNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGT
L
FSTEEEPSKLLVQNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYF
IARKIRKKRLENRKSVVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

[illegible][illegible]

FIGURE 244

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 245

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCG
GAGCCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCG
GCAGCCGGGAGCCATGCGACCCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCC
TGCTGCTCCTGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAG
CAAAAGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGG
GCCAGCAGGAGTGCTTGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTG
GGATCCCAGGTCGGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAG
GAGTCCTGGACACCCAACCTACAAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCT
TGGGAAAATTGCGGAGTGACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGT
TCAGTGGCTCACTTCGGCTAAAATGCAGAAATGCATGCTGTCAGCGTTGGTATTTACATTC
AATGGAGCTGAATGTTTCAAGACCTCTTCCCATTGAAGCTATAATTTATTTGGACCAAGGAAG
CCCTGAAATGAATTCACCAATTAATATTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAG
GAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGGTGGCACTTGTTTCAGATTACCCAAAA
GGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATCATTATTGAAGAACTACCAAAATA
AATGCTTTAATTTTCATTTGCTACCTCTTTTTTTTATTATGCCTTGGAATGGTTCACTTAAAT
GACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGA
CCAAAGTGATTTTCACACTGTTTTTAAATCTAGCATTATTCATTTTGCTTCAATCAAAAGT
GGTTTCAATATTTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTA
TAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTTAAAAAATA
TAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTTATATCTGTTAAAT
AAAAATTATTTCCAACA

FIGURE 246

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393
><subunit 1 of 1, 243 aa, 1 stop
><MW: 26266, pI: 8.43, NX(S/T): 1
MRPQGPAASPQRLRGLLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPA
PGRDGSPGANVIPGTPGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEATIIYLDQGSP
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRITIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

for the 249/249